

Stoyles, Chloe Louisa (2016) *An investigation into the folding and assembly of engineered antibodies and novel antibody formats*. PhD thesis.

<http://theses.gla.ac.uk/7675/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

An Investigation into the Folding and Assembly of Engineered Antibodies and Novel Antibody Formats



UNIVERSITY
of
GLASGOW

Chloe Louisa Stoye BSc (Hons)

Supervisors: Professor Neil Bulleid and Dr. Katharine Cain

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Institute of Molecular Cell and Systems Biology

University of Glasgow

October 2016

Abstract

Monoclonal antibodies and novel antibody formats are currently one of the principal therapeutics in the biopharmaceutical industry worldwide and are widely used in the treatment of autoimmune diseases and cancer. It is for this reason that the productivity and quality of antibody production requires improvement; specifically investigations into the engineering of antibodies and any issues that may arise from the production of these therapeutics. The work presented in this thesis describes an investigation into the folding and assembly of seven antibodies plus the novel antibody format FabFv.

IgG is comprised of two identical HCs and two identical LCs. The folding process of immunoglobulin is controlled by the CH1 domain within the HC. The CH1 domain remains in a disordered state and is sequestered by BiP in the endoplasmic reticulum. Upon the addition of a folded CL domain, BiP is displaced, the CH1 domain is able to fold and the complete IgG protein can then be secreted from the cell. The results presented in this thesis however, have outlined an additional mechanism for the folding of the CH1 domain. We have shown that the CH1 domain is able to fold in the absence of LC resulting in the secretion of HC dimers in a VH dependent manner. The proposed mechanism for the secretion of HC dimers suggests that some VH domains can interact with each other in order to bring the CH1 domains in close proximity to enable folding to occur. As HC dimer secretion is a hindrance in antibody production, this result has highlighted an engineering target to improve antibody yield. Examination of the folding of IgG4 with the variable region A33 has revealed the inability to secrete LC dimers, cleavage of the HC during expression and secretion of HC dimers in the Fab, FabFv and full length forms. The attributes described have also been shown to be variable region dependent. This has introduced a new concept that the variable domain is important in determining the expression and secretion of antibodies and their individual chains.

Pulse chase and 2D gel electrophoresis analysis of the novel antibody format FabFv has revealed that the folding and expression of the LC and HC causes multimeric species of FabFv to be secreted, as opposed to the monomeric form which is the desired therapeutic. Our hypothesis is that this process occurs via a LC dependent mechanism. The proposed hypothesis suggests that further engineering to the LC could diminish the formation and secretion of FabFv multimers. The results from these investigations can be applied to increase the productivity of therapeutics and increase the biological understanding of the domain interactions of IgG during folding, assembly and secretion.

Table of Contents

Abstract.....	ii
Table of Contents.....	iii
List of Tables	vii
List of Figures	viii
Acknowledgement	x
Author's Declaration	xi
Definitions/Abbreviations.....	xii
Chapter 1 : Introduction	0
1.1 Overview	1
1.2 Immunoglobulins	1
1.2.1 Role and development of antibodies in the immune response.....	1
1.2.2 B cell development and variable region generation.....	2
1.2.3 Different classes of Immunoglobulins.....	4
1.2.4 IgG structure and function.....	7
1.3 Translation, assembly and secretion of IgG	13
1.3.1 Quality control in the endoplasmic reticulum	13
1.3.2 Disulphide bond formation	17
1.3.3 Antibody folding and assembly in the endoplasmic reticulum.....	19
1.3.4 LC and HC dimer secretion.....	23
1.4 Antibodies and novel antibody formats used as therapeutics	24
1.4.1 Humanisation of antibodies	24
1.4.2 Generation of therapeutics using CHO cells	26
1.4.3 Antibodies and novel antibody formats developed for therapeutic use.....	27
1.4.4 FabFv structure and function.....	31
1.5 Summary and project aims	31
Chapter 2 : Materials and Methods.....	33

2.1	Cell Culture.....	34
2.1.1	Cell maintenance	34
2.1.2	Transient transfection of DNA constructs.....	34
2.2	DNA Methods.....	35
2.2.1	Site-directed mutagenesis	35
2.2.2	Sub cloning of DNA fragments into vectors by PCR or restriction digests.....	35
2.2.3	Mini and Midi plasmid DNA purification from transformed <i>E. coli</i>	41
2.3	Protein Methods	41
2.3.1	Western blot analysis.....	41
2.3.2	Pulse chase assay	42
2.3.3	2D gel electrophoresis	43
2.3.4	Semi-permeabilised (SP) cell assay	43
2.3.5	Trichloroacetic acid (TCA) precipitaion of the cell medium.....	44
2.3.6	Silver stain	44
2.3.7	Preparation of protein samples for mass spectrometry using a trypsin digest.....	45
2.3.8	Protein digests using Endo H and PNGase F	46
2.4	Protein expression and purification.....	46
2.4.1	Expression and Induction of Protein using Isopropyl β -D-1-thiogalactopyranoside (IPTG) 46	
2.4.2	Purification of the CH1 Domain from Inclusion bodies	47
2.4.3	Purification of CL and A33 variable region domains.....	48
2.4.4	Circular Dichroism (CD)	48
Chapter 3	: An investigation into the cleavage of the HC of IgG4 with the variable domain A33.	50
3.1	Introduction	51
3.2	Aim	54
3.3	Identification of a cleavage product of A33 IgG4	54
3.4	Classification of the cleavage product of A33 IgG4	57
3.5	Investigation of IgG4 mutants to address why the cleavage product is not seen with IgG1	

3.6	Cleavage of the HC A33 IgG4 only occurs with this construct	62
3.7	Discussion.....	64
Chapter 4 : An investigation into the secretion of HC dimers in a LC independent manner		67
4.1	Introduction	68
4.2	Aim	69
4.3	HC dimer secretion occurs in the presence and absence of LC	69
4.4	Confirmation of HC dimer secretion	71
4.5	Sequence variation of the VH domain affects the level of secretion of the A33 HC dimer 77	
4.6	HC dimer secretion is variable region dependent	80
4.7	LC dimer secretion is variable region dependent	84
4.8	Generation and Secretion of Fab Fragment of A33 IgG4 HC	86
4.9	The A33 variable domain folds the CH1 domain by means of an alternative mechanism which is used by the CL domain	88
4.9.1	Purification of individual immunoglobulin domains.....	90
4.9.2	CD of immunoglobulin domains in isolation.....	96
4.9.3	Mixing individual CL and CH1 domains introduces structure into the CH1 domain which is not seen when the CH1 and A33 VH domains are combined.	98
4.10	Discussion.....	101
Chapter 5 : An investigation of the folding, assembly and secretion of FabFvs.....		104
5.1	Introduction	105
5.2	Aim	108
5.3	Identification of FabFv folding intermediates.....	108
5.4	LC FabFv expressed alone results in multimer formation and secretion	113
5.5	HC expressed alone results in dimer formation and secretion	115
5.6	Summary of initial intermediates formed and secreted during FabFv assembly	117
5.7	Discussion.....	119
Chapter 6 : Discussion.....		124
6.1	Overview	125

6.2	How and why is the HC A33 IgG4 being cleaved?.....	125
6.3	Applications and causes of HC dimer secretion.....	127
6.4	Engineering of FabFvs to avoid multimer formation	132
6.5	Conclusions	133
Chapter 7 : Appendices.....		135
7.1	Appendix 1	135
7.2	Appendix 2	137
7.3	Appendix 3	138
7.4	Appendix 4	139
List of References.....		140

List of Tables

Table 1: Table of constructs used and generated in this thesis.....	38
Table 2: Table of antibodies with corresponding catalogue numbers used in this thesis.....	41
Table 3: Table of proteins produced after expression in CHO cells from constructs generated....	135
Table 4: Table of primers used for mutagenesis and sub-cloning protocols.....	137

List of Figures

Figure 1.1: Variable Region Gene Rearrangements of the HC.....	3
Figure 1.2: Schematic representations of each antibody class.....	6
Figure 1.3: Schematic representation of a Fab arm binding to its antigen via the hypervariable region using the six CDRs from both variable domains.	8
Figure 1.4: Crystal structure of IgG.	9
Figure 1.5: Interchain disulphide bond arrangements of each IgG isotype.....	11
Figure 1.6: Schematic representation of the Calnexin/Calreticulin quality control cycle in the ER.	15
Figure 1.7: Disulphide bond formation in the endoplasmic reticulum.....	18
Figure 1.8: Antibody folding and assembly in the endoplasmic reticulum.	21
Figure 1.9: Schematic representation of a selection of various antibody formats utilised as therapeutics.....	29
Figure 2.1: Plasmid maps of the vectors pMhg4PFL, pet-28a and pcDNA 3.1 used to generate each construct.	40
Figure 3.1: Schematic structure of IgG and each domain.....	53
Figure 3.2: Pulse chase analysis of IgG1 and IgG4 with the variable region A33.....	56
Figure 3.3: Identification and classification of the cleavage product of A33 IgG4.	59
Figure 3.4: Comparisons of IgG1 and IgG4 constructs to identify evidence of the cleavage product.	61
Figure 3.5: Screen of seven IgG4 HCs for presence of the cleavage product.	63
Figure 4.1: HC dimer secretion occurs in the absence and presence of LC.	70
Figure 4.2: Assembly and secretion of A33 IgG4 analysing the composition of various antibody intermediates.....	73
Figure 4.3: Confirmation of HC dimer secretion via mutagenesis and glycosylation studies.....	76
Figure 4.4: Contrast of variable region engineering via the humanisation process.	79
Figure 4.5: HC dimer secretion of seven different HCs of IgG4.	81
Figure 4.6: Sequence alignments and VDJ assignments of the seven VH domains investigated.	83
Figure 4.7: LC assembly and secretion of 497 and A33 in the absence of HC.	85
Figure 4.8: Generation and evaluation of the assembly and secretion of the truncated A33 IgG4 HC.....	87
Figure 4.9: Two potential mechanisms as to how the VH domain is folding the CH1 domain.....	89
Figure 4.10: Solubility test of each individual immunoglobulin domain expressed.	91
Figure 4.11: Purification of CH1 domain from inclusion bodies.	91
Figure 4.12: Purification and analysis of isolated A33 VH domain	93
Figure 4.13: Purification and analysis of isolated CL domain.	95

Figure 4.14: Far UV CD Spectra of CH1 domain (A), CL domain (B) and A33 VH domain (C).	97
Figure 4.15: Far UV CD Spectra of CH1 domain mixed with either CL domain or A33 VH domain.	100
Figure 5.1: Schematic of the structure of FabFv.....	107
Figure 5.2: Assembly and secretion of FabFv 497 over time.	109
Figure 5.3: 2D gel analysis of FabFv folding intermediates and secreted products.	112
Figure 5.4: Assembly and secretion of LC FabFv.....	114
Figure 5.5: Assembly and secretion of FabFv HC.	116
Figure 5.6: Summary of FabFv assembly and secretion.....	118
Figure 5.7: Model of multimer assembly of FabFvs.....	121
Figure 6.1: Hypothesis for the folding, assembly and secretion of HC dimers.....	130
Figure 7.1: Appendix 2: Confirmation of separation of the endoplasmic reticulum from the cytosol of the cell.	138
Figure 7.2 Calibration data using protein standards for Superdex 200 10 300 gel filtration column.	139

Acknowledgement

I would like to thank UCB Pharma and BBSRC for funding my PhD, I have thoroughly enjoyed the opportunity to participate in a collaborative project.

I would like to express gratitude to my incredibly supportive supervisor Professor Neil Bulleid (University of Glasgow) for all his help and guidance throughout my PhD. I gained a lot of knowledge and experience working with Neil who always took the time to teach and discuss my project with me and encouraged me to include my own thoughts and ideas to develop my skills as an independent scientist.

I would also like to thank my supervisor at UCB Pharma, Dr Katharine Cain for all her support and communications during this collaboration. I wish to also say a big thank you to Dr. David Humphreys, Dr Paul Stephens and the staff at UCB for all their helpful ideas and enthusiastic discussions of my PhD.

I would like to thank several people who have helped me with specialised techniques which were required during my investigations. Firstly, thanks to Dr. William Mullen for providing the mass spectrometry data and protocol which led to identification of the cleavage product. Secondly, thank you to Dr Sharon Kelly for carrying out circular dichroism experiments and for all her help in taking the time to go through the data with me. Finally, thanks to Professor William Cushley for his role as my advisor and for aiding me with identification of VDJ genes.

All the members of the Bulleid lab have been incredibly helpful and encouraging during my PhD and I would like to thank everyone involved; both past and present members who include Fiona Chalmers, Dr Philip Robinson, Dr Greg Poet, Dr Rachel Martin, Donna McGow, Esraa Haji, Tomasz Szmaja, Lorna Mitchell, Dr Zhenbo Cao and Dr Marcel Van Lith. In particular I would like to say a special thank you to my friend and colleague Marie Anne Pringle who taught me molecular biology methods and provided me with much needed “sanity-maintenance” over the past four years. I would also like to thank Dr Ojore Oka for assisting me with protein purification methods and donating his time and wisdom in the process.

Finally, I would like to say a special thank you to my family for all their love and reassurance throughout my life and particularly over the past four years. My father Andrew, my mother Jacci and my brothers Sam and Jake are all wonderful people who always support me to achieve my goals and I would not have been able to do this without them.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Chloe Stoye

October 2016

Definitions/Abbreviations

α 1AT	Alpha 1 anti trypsin
ADCC	Antibody-dependent cell mediated cytotoxicity
ATF6	Activating transcription factor 6
BiP	Immunoglobulin heavy-chain binding protein
BME	β -Mercaptoethanol
CD	Circular dichroism
CDC	Complement-dependent cytotoxicity
CDR	Complementary determining region
CH1	First constant domain (HC)
CH2	Second constant domain (HC)
CH3	Third constant domain (HC)
CH4	Fourth constant domain (HC)
CHO	Chinese hamster ovary
CHO-S	Chinese hamster ovary suspension
CL	Constant domain (LC)
D	Diversity
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide

DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
ER	Endoplasmic reticulum
ERAD	ER associated degradation
Ero1	Endoplasmic reticulum oxidoreductin 1
Fab	Fragment antibody binding
Fc	Fragment crystallisable
Fv	Fragment variable
GA	Golgi apparatus
GalNAc	N-acetylgalactosamine
GdHCl	Guanidine hydrochloride
GF	Gel filtration
GFP	Green fluorescent protein
GSH	Glutathione
GSSG	Glutathione disulphide
h	Hour
HC	Heavy chain

HC ₂	Heavy chain dimer
Heptamer	Seven nucleotides
Hsp	Heat shock protein
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPTG	Isopropylthio-β-galactoside
IRE1	Inositol requiring kinase
J	Joining
kDa	Kilodalton
L	Litre
LC	Light chain
LC ₂	Light chain dimer
LC ₃	Light chain trimer
M	Molar
MHC	Major histocompatibility complex

min	Minute
MMP	Metalloproteinases
NEM	N-ethylmaleimide
Nonamer	Nine nucleotides
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDI	Protein disulphide isomerase
PEI	Polyethylnimine
PERK	Protein kinase-like endoplasmic reticulum kinase
PPIase	Peptidyl prolyl <i>cis/trans</i> isomerase
Pre-BCR	Pre-B cell receptor
RA	Rheumatoid arthritis
scFv	Single chain Fv
SDS	Sodium dodecyl sulfate
SF	Soluble fraction
SLC	Surrogate light chain
SP	Semi-permeabilised

TBS	Tris buffered saline
TBST	Tris buffered saline with tween
TCA	Trichloroacetic acid
UPR	Unfolded protein response
UV	Ultra violet
V	Variable
VH	Variable domain (HC)
VL	Variable domain (LC)
v/v	Volume/volume
WCL	Whole cell lysate
WT	Wild type
w/v	Weight/volume
XPB1	Xbox binding protein 1

Chapter 1 : Introduction

1.1 Overview

Therapeutic monoclonal antibodies are the leading product in terms of profitability in biopharmaceuticals today. In 2013 antibodies used as therapeutics generated a net worth of \$140 billion, making the development and quality control of the products generated a vital area of current research (Walsh, 2014). However, the general cost of manufacturing a biopharmaceutical in the United States is approximately \$45 a day compared to \$2 a day for small chemicals (drugs), therefore developments in the production of these therapeutics to increase the economic potential is required to lower the manufacturing costs (Blackstone & Joseph, 2013). Treatment using therapeutic antibodies requires a high dose over a long time period, increasing the demand for an efficient production and high quality of product (Rodrigues, Costa et al., 2010). Several bottlenecks in the production of therapeutic antibodies have been identified; these include insertion of DNA into the CHO (Chinese Hamster Ovary) genome, transcription and translation events, protein folding and assembly, culture environments, redox status and protein overexpression, all of which contribute to the yield of antibody production (Mason, Sweeney et al., 2012, Rita Costa, Elisa Rodrigues et al., 2010). The successful development of therapeutic antibodies has led to the generation of novel antibody formats which combine the knowledge and technology gained from monoclonal antibody production to generate products with more tailored functions. The focus of this thesis is to investigate the bottleneck of antibody folding and assembly with the goal that a greater understanding of the process will lead to an opportunity to make the assembly and secretion of antibodies more efficient therefore increasing productivity.

1.2 Immunoglobulins

1.2.1 Role and development of antibodies in the immune response

The natural immune response of humans involves the removal of harmful pathogens to combat disease and keep the host healthy. There are two general types of immunity which exist to protect the human body which are adaptive and innate immunity. Innate immunity involves cells called phagocytes which are able to recognise and destroy harmful micro-organisms and pathogens by engulfing and eliminating them from the body and is described as a more generalised approach in defending against potential infections. However, adaptive immunity is a more specific response involving lymphocytes which are able to recognise, destroy and remember particular pathogens so that the immune system can attack the pathogen more readily if it is faced with it again (Hoebe, Janssen et al., 2004). The lymphocyte cells involved in adaptive immunity can be separated into two different types; B cells and T cells. T cells either aid B cells and other phagocytes in a co-operative fashion or have a more direct role whereby they recognise

cells which contain a virus via the major histocompatibility complex (MHC) proteins and subsequently destroy them. B cells are responsible for antibody production and secretion. Antibodies are highly specific multi-domain glycoproteins that recognise epitopes on the antigen target. The antibody-antigen complex is able to recruit T cells and other proteins in the immune response such as the proteins involved in the complement system via the Fc fragment (for further explanation of antibody structure see section 1.2.4) (Dunkelberger & Song, 2010b). The complement system is made up of approximately 30 proteins which evoke a cascade mechanism which ultimately destroys the pathogen via a communication of both the adaptive and innate immunity responses (Dunkelberger & Song, 2010a, Sarma & Ward, 2011).

1.2.2 B cell development and variable region generation

The ability of the immune system to “remember” pathogens and produce antibodies of such high specificity involves the cooperation of protein interactions and genetic re-modelling. The development of B cells and the generation of the variable domain is a complex process that ultimately produces a highly specific protein which can protect the host from the pathogens invading it.

B cell development involves a three stage process to reach an immature antibody-producing cell (reviewed in (Pieper, Grimbacher et al., 2013). The first stage involves gene rearrangements to produce a new heavy chain (HC); the cells in this first stage reside in the bone marrow and are called pro B cells. Variable region generation of the HC involves the fusion of three different gene segments (Figure 1.1). First, a single D (diversity) and a single J (joining) gene fuse together, each gene segment being randomly selected. A single V (variable) gene is then fused to the DJ segment, which now becomes a combination of VDJ encoding for the variable domain of the HC. The correct order of the fusion of each gene segment is governed by the orientation of the DNA helix and number of nucleotides within them to ensure each gene segment is formed correctly; known as the 12/23 rule. The 12/23 rule determines the recombination signal sequence and is comprised of a heptamer (seven nucleotides), a nonamer (nine nucleotides) or both spaced by either 12 or 23 nucleotides (Schatz & Ji, 2011). The heptamer-nonamers determine the orientation of the DNA helix as only when they are orientated in opposite directions and are separated by 12 or 23 nucleotides does joining of the DNA segments occur. This rule ensures that the LCs only make V-J combinations and the HCs only make V-(DJ) combinations; failure of this process results in immune deficiency diseases (Moshous, Callebaut et al., 2001). The VDJ gene then fuses with the constant gene of IgM, or sometimes IgD, enabling the transcription and translation of a functional HC.

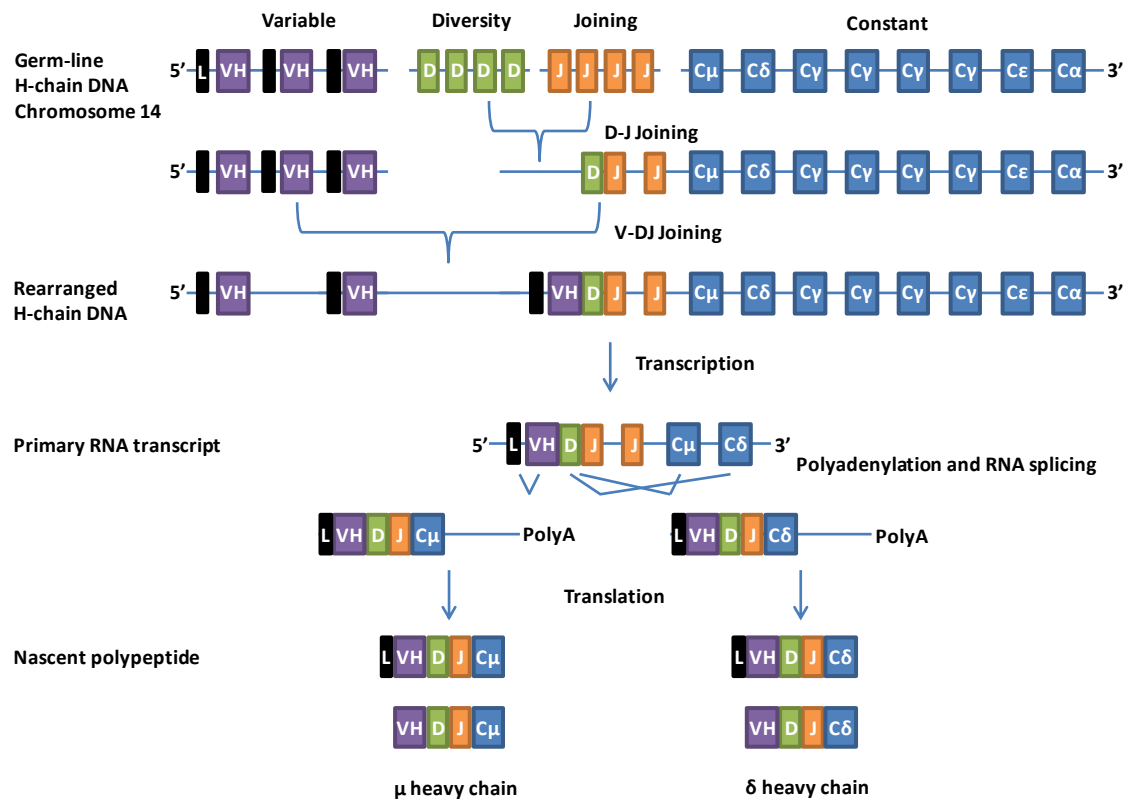


Figure 1.1: Variable Region Gene Rearrangements of the HC.

The VH domain of the HC is organised by the selection of random gene segments which are spliced together. Firstly, diversity (D) and joining (J) gene segments are selected randomly and spliced together forming a DJ exon. Following this, a random variable (V) gene segment then joins with DJ to form VDJ which encodes for the variable region of the HC. The VDJ is then joined with either a Cμ or Cδ constant gene segment to produce a functional HC protein.

The next process in B cell development is split into two stages; early and late pre-B cell development (late pre B cell development is recognised by the cell proliferating into an immature B cell). The difference between the two is dependent on the type of LC expressed. In early B-cell development, a surrogate light chain (SLC) associates with the newly synthesised μ HC in order to aid in folding the C μ 1 domain of the HC for expression on the cell surface. The SLC is made up of two separate proteins which associate together; the Vpre-B protein which represents the VL domain of LCs and the λ 5 protein which represents the CL domain of LCs and is covalently linked to the C μ 1 domain of the HC via a disulphide bond (Roitt, Brostoff et al., 1998). The SLC shares homology with the LC of immunoglobulins but does not function as a LC post B cell development. The complex formed comprising of the SLC and μ HC is known as a pre-B cell receptor (pre-BCR). Once the pre-BCR has been expressed on the cell surface, signal transduction molecules also present on the cell surface interact with the pre-BCR to signal to the cell that a functional HC has been expressed. The next stage of development in producing a functional LC involves the rearrangement of V and J genes with either lambda (λ) or kappa (κ) constant genes to form a light chain (LC) to pair with the newly synthesised HC. The late pre-B cells, however, cannot progress to this stage until a protective screen is performed by the pre-B cell antigen receptor complex which inhibits further HC gene rearrangement and initiates LC rearrangement. Once the HC has paired with the LC the newly synthesised IgM antibody is expressed on the cell surface where it is known as a B-cell receptor capable of binding to an antigen. The proliferation and further development of the cell will continue with the commitment of binding to its specific antigen (Mårtensson & Ceredig, 2000). The cell then migrates to the spleen as an immature B cell where it can further differentiate into a B cell.

The combination of random VDJ genes with random LC and HCs plus high mutation processes to generate a higher specificity of binding give rise to the potential of 5×10^{13} different possibilities of variable regions. This number demonstrates how the immune system is able to achieve a defence mechanism against such a large range of potential pathogens.

1.2.3 Different classes of Immunoglobulins

There are five classes of immunoglobulin (A, D, E, G and M) which differ in their constant regions and are produced by B cells (Figure 1.2). As described earlier, B cell development for the synthesis of new HCs combines a random VDJ sequence with the constant genes of IgM or IgD. However, in order to select the correct isotype for the function of the type of immune response required, different immunoglobulin classes are produced. This process of class switching occurs in a T-cell dependent fashion whereby different chemical signals will cause the genetic class-switch of a B cell to produce a different isotype (Peng, Szabo et al., 2002). All 5 genes for each separate

antibody class are located upstream of IgM constant gene on chromosome 14 (Figure 1.1). Prior to each gene is a class switching sequence that is activated by different signalling molecules which bind to the correct antibody class which is appropriate to the type of immune response required (Roitt et al., 1998). This enables the production of a new isotype with the same variable region previously generated to attack the antigen with the correct action to deal with the particular immune reaction.

The structure and function of each immunoglobulin class are highly variable (reviewed in (Schroeder & Cavacini, 2010)). All HCs consist of one variable and three constant domains (except IgM and IgE which have a fourth constant domain), with each class and isotype differing in the sequence of the constant regions allowing them to bind to their corresponding Fc receptors. The difference in quaternary structure of each antibody class is outlined in Figure 1.2. Secreted IgM exists as a pentameric structure and is involved with the primary immune response, which is similar to IgD with a “Y” structure that is involved in antigen recognition. The pentameric structure of IgM is stabilised via disulphide bonds at the C-terminal part of the protein and the J chain, which is a small polypeptide linker produced in mucosal and glandular plasma cells (Johansen, Braathen et al., 2000). IgA has two isotypes (IgA1 and IgA2) which exist as a dimer and are involved in mucosal immunity attacking pathogens which are inhaled or ingested. IgA exists as a dimer, where the C-terminus of the protein forms a disulphide bond with a J chain. IgA also contains a secretory component derived from the polymeric immunoglobulin receptor in endosomes that is required for its secretion (Phalipon, Cardona et al., 2002). IgE has a “Y” structure and is involved in the allergic response; it does not contain a hinge region and has an additional interchain disulphide between its CH2 domains, therefore making its CH3 and CH4 domains more homologous to the CH2 and CH3 domains of IgG. IgG consists of four isotypes (IgG1, IgG2, IgG3 and IgG4) which have a “Y” structure and are involved in a secondary immune response, usually IgG is the first class switch from IgM after the invasion of a pathogen and is the most abundant antibody in the blood. All classes of immunoglobulin are glycosylated within the CH2 domain; this glycosylation is important in the overall stability of the protein. The complexity and necessity of glycosylation varies with each class of antibody and is predominately involved in the interactions with its corresponding Fc receptor binding via interfaces of the Fc receptors hydrophobic core. Glycosylation is essential for Fc receptor binding to IgG, for production of IgD and is a necessary overall contributor to the function of each isotype (Liu, Bulseco et al., 2006).

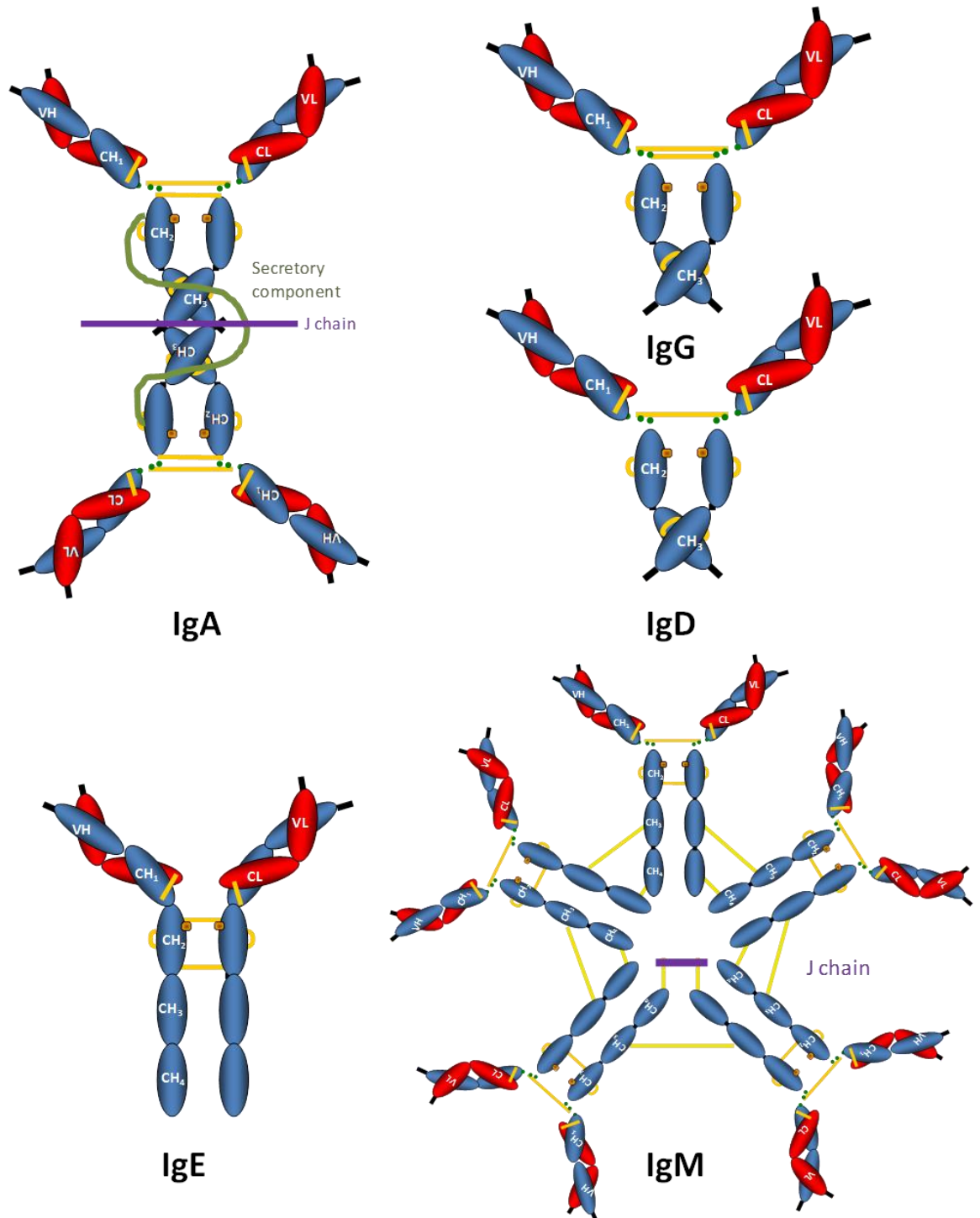


Figure 1.2: Schematic representations of each antibody class.

There are five different classes of immunoglobulin (A, G, D, E, and M). Each class differ in their structure and function in the immune system. Each antibody structure is comprised of two identical LCs (red) and two identical HCs (blue) held together via disulphide bonds (yellow). Within each structure the CH₁ and CH₂ domains are joined by a hinge region (dark green) with the exception of IgE. IgA and IgM have an additional element of a J chain (purple) to their structure to exist as a dimer (IgA) or pentamer (IgM). IgA also has a secretory component (light green) which is required for the protein to be secreted. All immunoglobulin are glycosylated within their CH₂ domains (orange) to aid in Fc receptor binding.

1.2.4 IgG structure and function

IgG is a symmetrical multi-domain protein comprised of two identical LCs and two identical HCs held together via interchain disulphide bonds. The HC is made up of four domains; three constant (CH (1-3)) and one variable (VH) domain, whereas the LC is half the size of the HC comprising of two domains; one constant (CL) and one variable (VL). The VH and VL domains together form the hypervariable region at the N-terminus of the molecule and are responsible for antigen recognition. This hypervariable region can be further divided on each domain into three complementary determining regions (CDRs), together all six CDRs combined form the antigen recognition site (Figure of 1.3). CDR3 on the HC is the most variable of all of the CDRs consisting of a mixture of gene splicing between V D and J segments of the antibody gene and interact with the antigen target. In comparison, CDRs 1 and 2 only comprise the V gene sequence. The overall structure of IgG can be divided into three sections; the Fab region, the hinge region and the Fc region which were first identified via proteolytic cleavage of pepsin and papain at the hinge region (Jefferis, Weston et al., 1968, Turner, Bennich et al., 1970). Each region of the antibody performs a different function in a cooperative manner; the Fab regions are responsible for recognition and binding to an antigen whilst the Fc region binds to Fc receptors to activate complement and the hinge region enables the flexibility of the protein to bind to both. The overall quaternary structure of the antibody gives a “Y” shape with the CH2 and CH3 domains dimerising at the base forming the Fc fragment, with the Fab arms consisting of the LC covalently bound to the VH and CH1 domains of the HC (Figure 1.4 A) (Deisenhofer, Colman et al., 1976a, Deisenhofer, Colman et al., 1976b). Each folded immunoglobulin domain has a similar overall secondary structure. Seven (constant) or nine (variable) β strands are arranged into two anti-parallel β pleated sheets which are internally stabilised by a hydrophobic core and an intrachain disulphide bond between each sheet (Figure 1.4 B) (Berg, Tymoczko et al., 2006). The sheets are joined together via flexible loops (CDRs in the variable domains) which require the isomerisation of a *cis* proline residue during their folding. The hinge region of the antibody also exists as a flexible loop structure, providing the immunoglobulin with its functional flexibility. Although the sequences of different classes and isotypes of immunoglobulins are different, the overall secondary structure of the immunoglobulin fold in each domain is highly conserved (Halaby, Poupon et al., 1999).

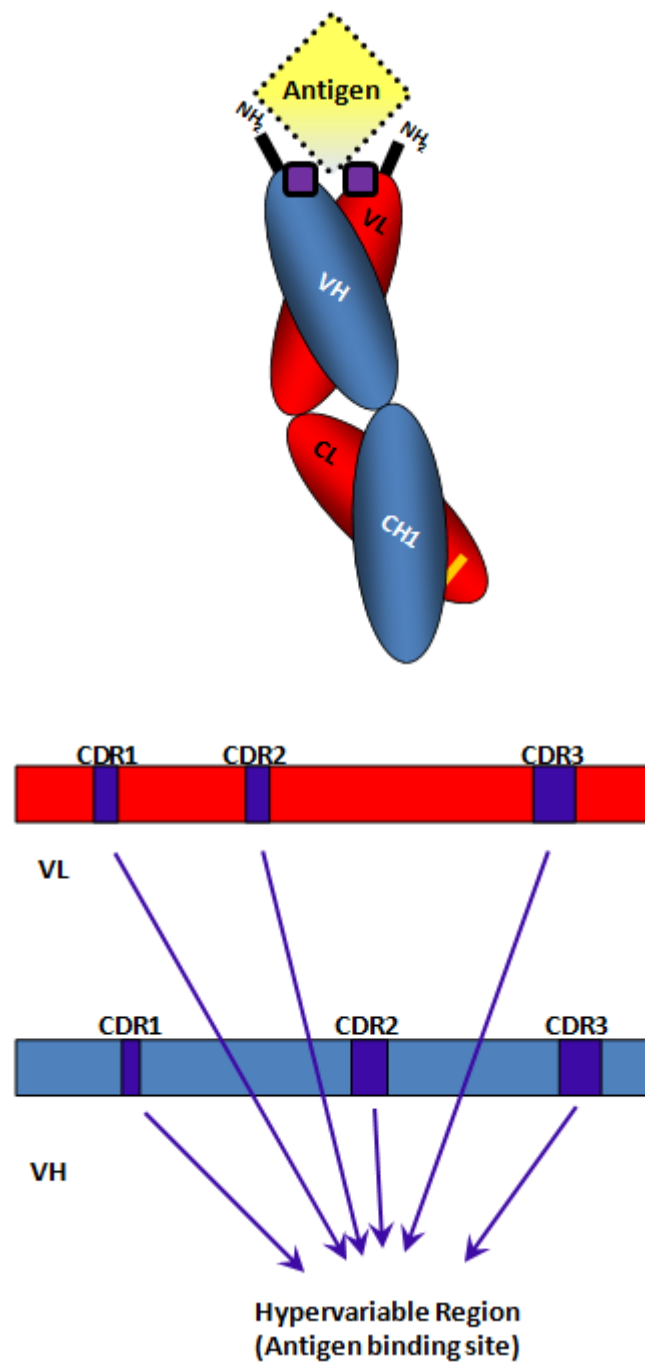


Figure 1.3: Schematic representation of a Fab arm binding to its antigen via the hypervariable region using the six CDRs from both variable domains.

Both the VL (red) and VH (blue) domains contain three CDR regions (purple); together which form the hypervariable region where the antibody binds to its antigen. The variable domain is the only immunoglobulin domain which is involved in antigen binding. The sequence between the CDRs known as the framework of the variable domains can have some interactions with the antigen but the main binding interface comes from the CDRs of both VL and VH domains.

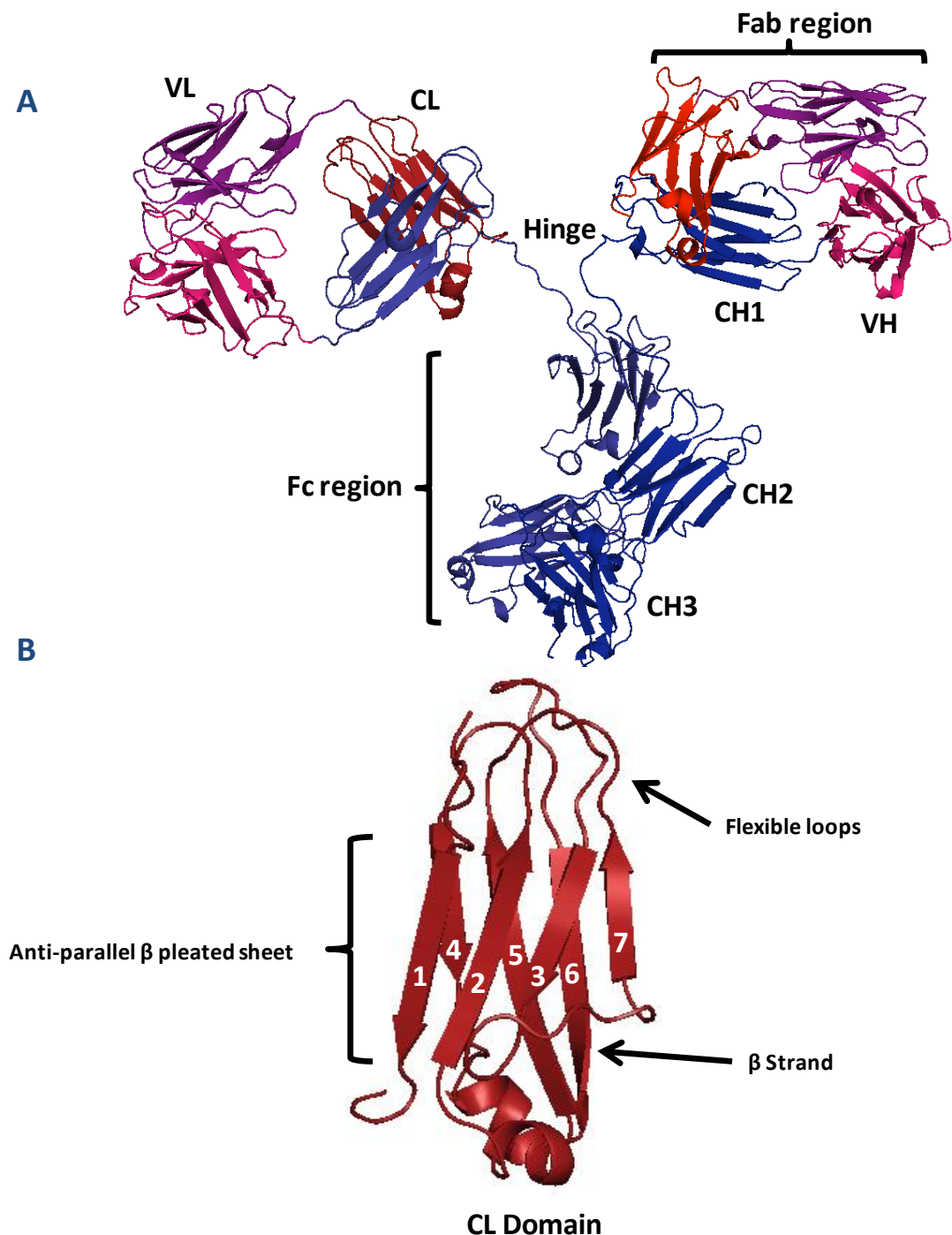


Figure 1.4: Crystal structure of IgG.

IgG is comprised of two identical LCs (red) and two identical HCs (blue) (A). The HC consists of four domains, one variable (VH) and three constant (CH1-3) whereas the LC is half the size and is only made up of two domains, one variable (VL) and one constant (CL). The LC and VH and CH1 domains of the HC make up the Fab region of the antibody which is involved in antigen binding. The CH2 and CH3 domains exist as a dimer, make up the Fc region of the antibody and are involved in Fc receptor binding. The Fc and Fab regions are held together via a flexible loop structure known as the hinge region, giving the antibody its function of binding to both its antigen and Fc receptor. Each immunoglobulin domain displays a similar overall secondary structure of two anti-parallel β pleated sheets held together via flexible loops (B). Images were generated using Pymol PDB code 1IGT and are of the antibody isotype IgG2 (Harris, Larson et al., 1997).

The four isotypes of IgG differ in their sequence of constant domains and their disulphide bond arrangement, which results in different effector functions for each antibody (reviewed in (Vidarsson, Dekkers et al., 2014) (Figure 1. 5). IgG1, IgG2, IgG3 and IgG4 are named as such due to their relative concentrations in the blood (Grey & Kunkel, 1964). The highest point of variability between each isotype is located in the hinge region of the antibody where interchain disulphide bonds link the two HCs. The interchain disulphide bonds vary with each isotype, the number of disulphide bonds within the hinge region of IgG is either two (IgG1 and IgG4), four (IgG2) or eleven (IgG3), also the interchain disulphide between the CH1 and CL domains differs as it is either within the hinge region for IgG1 or within the CH1 domain for IgG2, 3 and 4 (Liu & May, 2012). IgG1, being the most abundant IgG subclass has the function of recognising protein or peptide antigens either soluble or presented on membranes and is able to activate the complement system (Ferrante, Beard et al., 1990). IgG3 is also involved in activating complement, but to a lesser extent and has the main function of inducing effector functions in a pro-inflammatory response usually alongside IgG1. IgG2 mainly recognises polysaccharide antigen targets presented by bacteria, and has been implicated as the main IgG subclass involved in recognition of carbohydrate-based antigens (Siber, Schur et al., 1980). IgG4 is usually expressed after a prolonged exposure to a particular antigen and is therefore involved in the recognition of allergens, similar to IgE and is unable to activate the complement system like the other IgG isotypes (Jutel & Akdis, 2011). Therefore, the order of ability of IgG subclasses to activate the complement system via their CH2 domain follows IgG1>IgG3>IgG2.

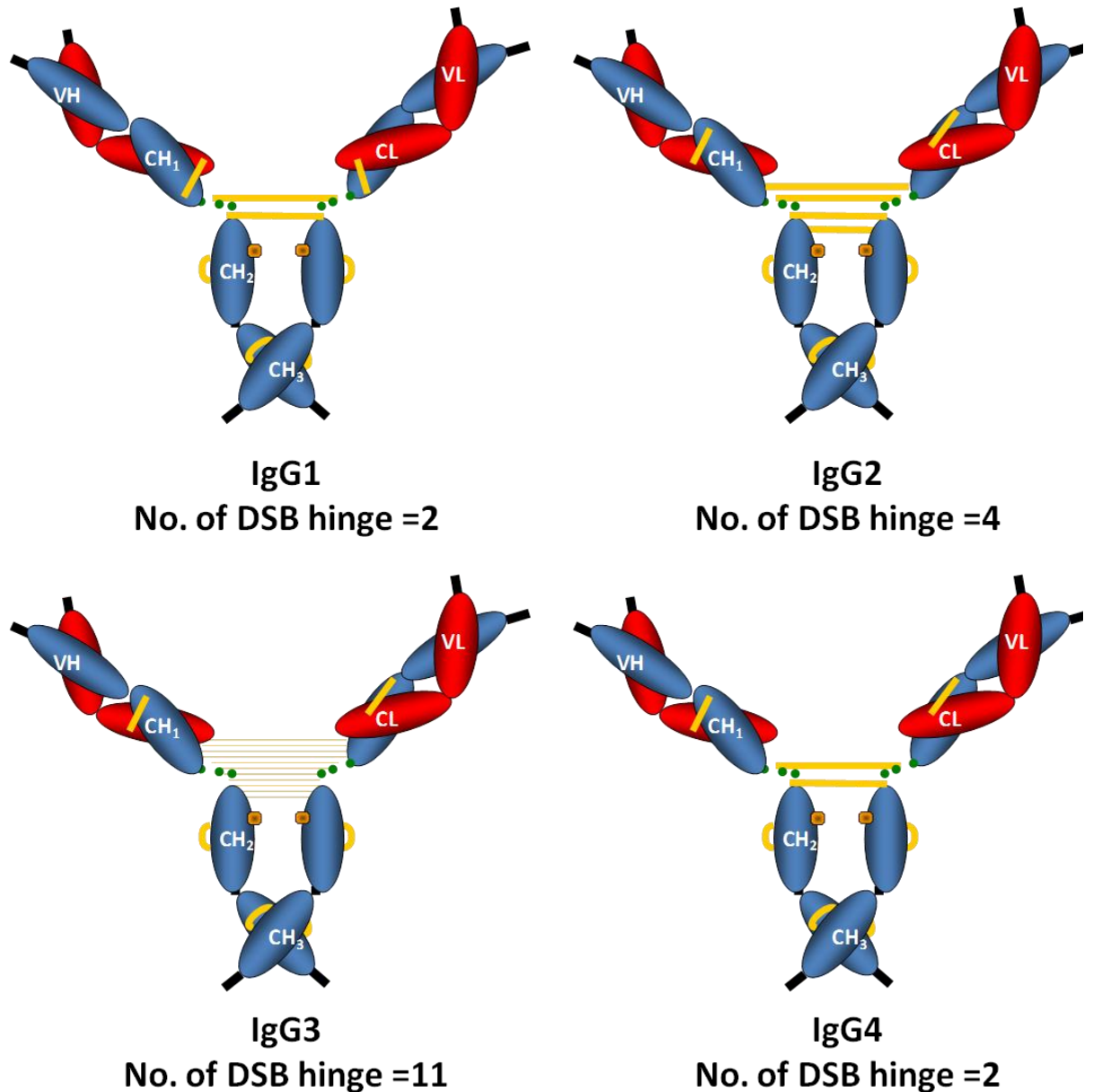


Figure 1.5: Interchain disulphide bond arrangements of each IgG isotype.

The four IgG isotypes differ in their sequence of their constant domains and disulphide bond arrangement. Each isotype is made up of two identical LCs (red) and two identical HCs (blue) held together via disulphide bonds (yellow). The hinge region of each isotype is the most variable in the number of disulphide bonds with IgG1 and IgG4 containing 2, IgG2 containing 4 and IgG3 containing 11. The interchain disulphide bond between the CL and CH1 domains also varies with each isotype. IgG1 contains a cysteine at position 220 within the hinge region to form an interchain disulphide with the CL domain whereas the other isotypes contain this cysteine at position 131 within the CH1 domain (amino acid numbering worked out using IgG1 and IgG4 with the variable domain A33).

IgG4 is a monovalent antibody, meaning it contains a single epitope or antigen target, however it has the ability to form bi-specific antibodies via a mechanism known as Fab arm exchange (or half molecule formation) after secretion. Half molecule formation is when the two antibody chains separate at the hinge region forming two halves of the same protein; these two HCs can then recombine to re-form a bi-specific antibody (Schuurman, Van Ree et al., 1999). The ability of two HCs to separate is caused by a substitution of a proline amino acid for a serine giving the sequence CPSC in IgG4, instead of CPPC in IgG1 within the hinge region (Aalberse & Schuurman, 2002). Half molecule exchange of IgG4 aids in reducing the level of inflammation in an immune response by inhibiting the formation of large immune complexes. Another aspect of how the isotypes are similar is their ability to activate complement-dependent cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC) (Clynes, Towers et al., 2000). This is achieved via the Fc region of the antibody by binding to proteins within the complement system (CDC) or recruiting cytotoxic cells to a particular pathogen (ADCC). As different isotypes display different abilities to activate complement, IgG1 is the isotype of choice for therapeutic antibodies as it can activate CDC and ADCC. If this activation is not required for the therapeutic action then IgG4 or IgG2 are the preferred isotypes to use (Parekh, Berger et al., 2012).

Investigations into increasing the thermal stability and avoiding the occurrence of half molecule formation of IgG4 have been characterised to further engineer the isotype for use as a therapeutic. For example, the CPPC sequence of IgG1 within the hinge region was introduced to the IgG4 sequence and once the immunoglobulin was expressed it was found to abolish the formation of half molecules (Angal, King et al., 1993). In addition to this, the differences within the hinge region and the interchain disulphide arrangement between the LC and HC were also investigated. A hinge swap between IgG1 and IgG4 which generated an IgG4 HC with the hinge sequence and interchain disulphide position to mimic that of IgG1 was generated. In addition, a range of potential interchain disulphide bond positions were investigated within the upper hinge to determine any consequences to thermal stability. The results of these investigations demonstrated an increase in thermal stability with the hinge sequence of IgG1 compared to the WT IgG4 and a flexibility of interchain disulphide bond formation as all positions investigated were able to form an interchain disulphide with the CL domain (Peters, Smales et al., 2012). This engineering of antibody isotypes to combine the positive attributes of different IgGs can be used to improve the quality of therapeutics which are generated.

1.3 Translation, assembly and secretion of IgG

1.3.1 Quality control in the endoplasmic reticulum

Ensuring a protein is folded correctly is essential for it to carry out its specific function. Therefore, the folding process and checkpoints for correct protein structure must be properly regulated. Within the ER, there are several folding chaperones which affect the rate of folding and aid the protein to reach its native structure. These chaperones can be grouped into different classes according to their function, which are; heat shock proteins (Hsp), peptidyl prolyl *cis/trans* isomerases (PPIases), oxidoreductases and glycan-binding proteins. Within these groups of folding factors lie several proteins responsible for post-translational modifications such as disulphide bond formation, glycosylation and proline isomerisation. However, there is also evidence of overlapping functionality of various proteins suggesting a greater level of complexity to the folding system. One example of this is protein disulphide isomerase (PDI) which is a member of the PDI family whose main function is disulphide bond formation and isomerisation (see section 1.3.2) but has also been shown to have other non-catalytic functions such as interacting with aggregated proteins, stabilizing them and targeting proteins for degradation (Braakman & Bulleid, 2011).

The overall process involved in ensuring proteins are correctly folded in the ER is known as quality control, and uses the glycosylation status of a protein to assess its folding status (Figure 1.6). The majority of proteins that enter the ER are initially glycosylated by oligosaccharyl transferase at an asparagine (Asn) residue located on the N-terminus of the growing polypeptide within the motif Asn-X-Ser/Thr (Grant, Welply et al., 1986). However, not all proteins with this motif are glycosylated. It was demonstrated that if the X residue is the amino acid proline or if this motif is located toward the C-terminal end of the protein then glycosylation is unlikely to occur and these non-glycosylated proteins would rely on the previously mentioned ER chaperones to aid in correct folding (Bause, 1983). This carbohydrate 'tag' which is added to the Asn residue consists of several oligosaccharides including mannose, glucose and N-acetyl glucosamine which are added as a pre-formed glycan core (Helenius & Aebi, 2004, Krambeck, Bennun et al., 2009). Trimming of glucose molecules by glucosidases enables the protein to enter the calnexin/calreticulin cycle where the protein can fold correctly and exit the ER. If this fails, UDP-glucose: glycoprotein glucosyltransferase (GT) re-glucosylates the protein to enable re-entry into the calnexin/calreticulin cycle in an attempt to fold the protein to its correct native state (Ritter & Helenius, 2000). However, if this is not achieved, the terminally misfolded protein follows the ER associated degradation (ERAD) pathway via the action of the enzyme EDEM which recognises the misfolded protein after the mannose molecule has been removed by α -1,2-mannosidase (Lipson,

Alalouf et al., 2008). The protein can then exit the ER for degradation in the cytosol following ubiquitination.

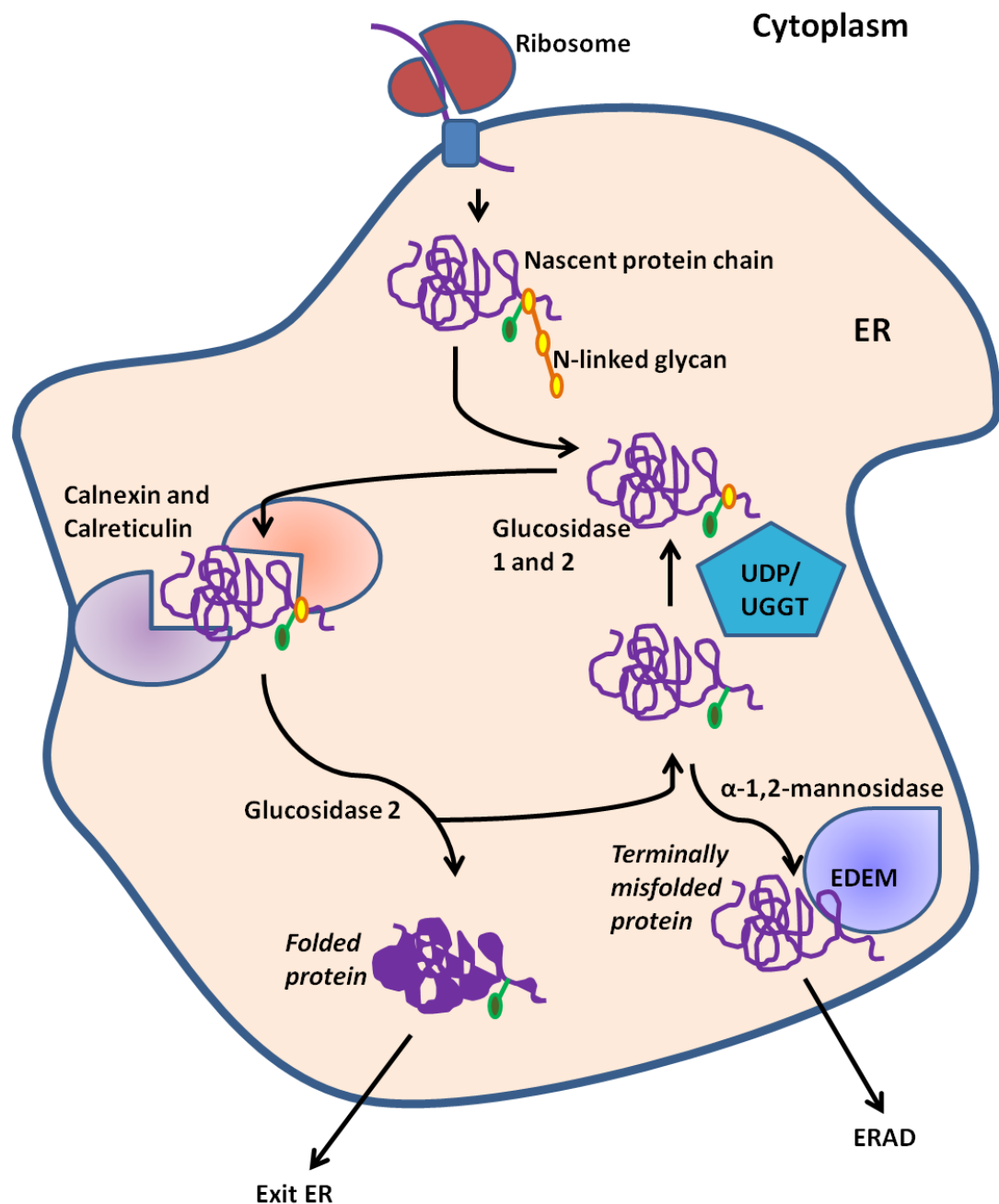


Figure 1.6: Schematic representation of the Calnexin/Calreticulin quality control cycle in the ER.

Proteins enter the ER as a nascent polypeptide chain and are subsequently glycosylated via an asparagine residue within an Asn-X-Ser/Thr motif. The yellow circles represent glucose and the green circles represent mannose. Glucosidase 1 and 2 removes two glucose molecules which enable the protein to interact with Calnexin and Calreticulin where the protein can fold with the aid of additional chaperones. Glucosidase 2 then removes the last glucose molecule enabling the folded protein to exit the ER. If the protein is misfolded, it is re-glucosylated by UDP/UGGT to enable re-entry to the Calnexin/Calreticulin cycle to fold correctly. If this fails and the protein is terminally misfolded, the mannose molecule is cleaved by α -1, 2-mannosidase and the protein is then targeted for degradation by ERAD via the action of the enzyme EDEM.

Conversely, if the correct structure is achieved, these proteins accumulate in COP II coated vesicles ready for transport to the Golgi apparatus; however, during this process several of the essential ER chaperones can be transported with them. In order to ensure these chaperones are returned to the ER there is a sequence (KDEL) at the C-terminus of the ER resident protein which is recognised by a complementary receptor within the Golgi apparatus (GA). The ER chaperones are then transported back to the ER in COP I coated vesicles completing the recycling process (Wilson, Lewis et al., 1993). Once the protein is in the GA, further post-translational modifications can occur. These include more complex glycosylation events either by modifying the N-linked glycans which were glycosylated in the ER or undergoing further glycosylation of O-linked glycans which is specific to the GA (Berg et al., 2006). Although O-linked glycosylation is specific to the GA, there have been some reported cases of initiation of O-linked glycosylation occurring in the ER or in the COP II coated vesicles that transport folded ER proteins to the GA (Röttger, White et al., 1998). O-linked glycosylation is named as such as it involves the addition of a N-acetylgalactosamine (GalNAc) to the hydroxyl (OH) group of a serine or threonine amino acid catalysed by GalNAc transferase (Lodish, Berk et al., 2000). During O-linked glycosylation, different sugars are added individually via specific enzymes according to the type of modification required instead of the pre-formed glycan core which is added to asparagine during N-linked glycosylation.

However, if the ER is put under stress such as a hypoxic environment, infection or glucose deprivation, its ability to fold proteins correctly, and deal with misfolded proteins can be hindered. The consequence of ER stress is the induction of a process known as the unfolded protein response (UPR) (reviewed in (Chakrabarti, Chen et al., 2011)). There are three main receptors within the ER membrane which initiate the UPR: inositol requiring kinase 1 (IRE1), protein kinase -like endoplasmic reticulum kinase (PERK) and activating transcription factor 6 (AFT6). These receptors are activated upon dissociation of an ER chaperone BiP. BiP, in its ADP bound state, has a high affinity for hydrophobic residues and, therefore, binds to misfolded proteins within the ER. The activation of these three receptors leads to a cascade of events which stall translation of new proteins in the hope to try and fold correctly the aggregated misfolded proteins already within the ER. If the ER can recover, BiP can re-associate with its corresponding receptors and continue with the translation of new proteins, however if recovery is not possible, this process ultimately results in apoptosis. The UPR illustrates how important correctly folded proteins are to the cell and highlights the tightly constructed quality control system the cell adopts.

1.3.2 Disulphide bond formation

Disulphide bonds are formed by a covalent linkage between the sulphur atoms of two cysteine residues. Any individual cysteine has the potential to form a disulphide bond, and these can be either intermolecular or intramolecular disulphides (Hatahet & Ruddock, 2007). The reaction of disulphide bond formation in the ER is catalysed by a group of enzymes called protein disulphide isomerases (PDI) of which there are approximately 17 family members. Collectively, this family of enzymes can oxidise, reduce or isomerise disulphide bonds within a substrate protein.

During folding of a protein, native and non-native disulphides are formed as various conformations are modelled until the correct structure is found. This process of disulphide formation occurs via the action of Ero1 proteins (amongst other possible pathways) which determines the oxidation state of PDI's enabling these enzymes to correctly orchestrate native disulphide formation (Figure 1.7) (Tavender & Bulleid, 2010). When an oxidised or reduced PDI family member starts the enzymatic interaction with its substrate a mixed disulphide is formed. This is when the PDI family member and its substrate have formed a covalent disulphide bond between two cysteine residues; from this conformation the reaction can continue to form an oxidised, reduced or isomerised disulphide (Hatahet & Ruddock, 2007). This electron transfer reaction is known as a thiol-disulphide exchange. However, the PDI family member must be re-oxidised in order to continue its function; this task is performed by Ero1 which oxidises PDI family members. Ero1 achieves this function by utilising molecular oxygen, to oxidise the PDI family member as well as producing hydrogen peroxide as a by-product.

There is still a lot unknown about disulphide formation, although it has been shown that this is the rate-limiting step in protein folding, and as the native disulphides of proteins are so different, the question of a possible higher level of specificity has emerged. For example, one member of the PDI family, P5, has been shown to strongly associate with BiP (a protein which recognises hydrophobic residues) and therefore could be involved in the disulphide formation of other proteins which associate with the chaperone compared to other PDI family members, this association suggests that each PDI family member may have some substrate specificity (Jessop, Watkins et al., 2009). Another example of substrate specificity of a PDI family member is ERp57 which has been shown to interact with glycosylated proteins as it interacts with the proteins calnexin and calreticulin (Elliott, Oliver et al., 1997).

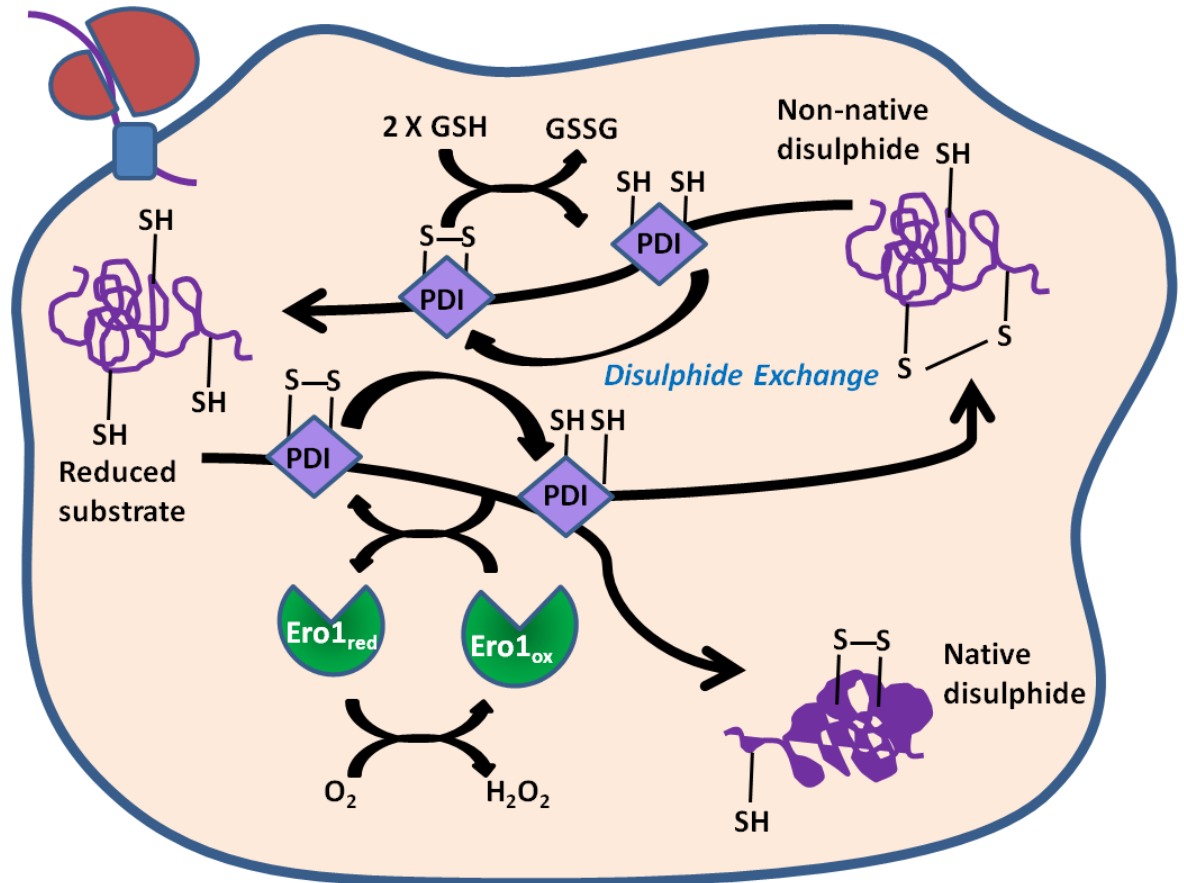


Figure 1.7: Disulphide bond formation in the endoplasmic reticulum.

Cysteine residues within the polypeptide chain contain free thiol (SH) groups which have the potential to form disulphide bonds. The proteins with these SH groups are in a reduced state. The reduced substrate enables action of PDI family members to introduce disulphide bonds into the protein oxidising the substrate in the process. This oxidation action leaves the PDI family member in a reduced state and therefore requires re-oxidation to act on another substrate. This action is performed by the ER chaperone Ero1. Ero1 is able to re-oxidise PDI family members by utilising molecular oxygen and producing hydrogen peroxide as a by-product. If a non-native disulphide is present in the protein the action by PDI is reversed. The reduction of PDI for this process is governed by a tripeptide molecule glutathione which forms a dimer via a disulphide and becomes oxidised after reducing a PDI family member. The process of reduction and oxidation of disulphides within the ER is known as disulphide exchange and can occur due to the equilibrium provided by glutathione and reactive oxygen species creating a redox environment where both oxidation and reduction can occur simultaneously (Oka, Yeoh et al., 2015).

1.3.3 Antibody folding and assembly in the endoplasmic reticulum

Antibody folding and assembly has been widely researched over the last 60 years. These investigations have characterised the order and rate of folding of each domain and is how several of the ER chaperones were first identified. One of the most important of these discoveries was the identification of the ER resident chaperone BiP (Bole, Hendershot et al., 1986). BiP has been shown to be essential in antibody folding as it strongly associates with the unfolded CH1 domain of the heavy chain resulting in the protein being retained in the ER therefore inhibiting the secretion of HC dimers with the other three domains on the HC already folded (Vanhove, Usherwood et al., 2001). Several proteins in the ER contain BiP-binding sequences within them, however, BiP association does not always occur if the rate of folding of the protein is high, this is the case for some of the other immunoglobulin domains (Hellman, Vanhove et al., 1999). For example, the CL domain of the LC contains BiP-binding sequences but as it folds rapidly upon entry into the ER with the first intrachain disulphide loop within the CL domain forming in less than one second, it has not been shown to interact with BiP (Bergman & Kuehl, 1979). The internal disulphide bridge within the CL domain has been shown to be pivotal in the folding of this domain as it restricts the potential interactions of residues within the domain enabling correct association and folding of the native protein to occur (Feige, Hagn et al., 2007). However, the LC can associate with BiP via the VL domain so that it is not secreted without its partner HC. The ability for BiP to bind to the VL domain varies between each LC due to the sequence differences between VL domains. BiP binding to the LC controls whether or not it is secreted in the absence of HC and contributes to the half-life of the protein before it is targeted for degradation (Davis, Khurana et al., 1999, Hellman et al., 1999, Lee, Brewer et al., 1999).

The folding of each domain of the HC follows a similar process as all domains have an almost identical secondary structure. All the constant domains of the HC require isomerisation of a conserved proline residue from a *trans* to a *cis* conformation at the flexible loops between the β strands of the protein. This reaction is catalysed by a PPlase and occurs before dimerization of the HCs within the Fc region and after association with the LC for the CH1 domain in the Fab region making it the final step in the folding of each domain (Lilie, Rudolph et al., 1995, Thies, Mayer et al., 1999). Each domain also has a strong hydrophobic core which initiates internalisation and therefore folding of each domain with the addition of an intrachain disulphide bond stabilising the internal structure of the immunoglobulin domains. The CH2 domain contains a glycosylation site which stabilises the HC dimer and is required for binding to Fc receptors, however, it has been shown that non-glycosylated CH2 domains still have the ability to associate (Feige, Walter et al., 2004). The CH3 domain initiates dimerization of the Fc fragment of the antibody; this is a strong association as it has been shown that the stabilising intrachain disulphide

within the CH3 domains is not a requirement for association (Thies, Talamo et al., 2002). Therefore, it can be said that dimerization of the HCs does not require the addition of post-translational modifications for paired domains to associate with each other.

The process of complete multi-domain folding of IgG is controlled by the unfolded CH1 domain of the HC (Figure 1.8). This domain also requires isomerisation of a proline residue; however it has been shown to remain unfolded unlike the other constant domains. The CH1 domain is unfolded and bound to BiP until the LC displaces BiP and associates with the CH1 domain. The displacement of BiP from the CH1 domain can be initiated *in vitro* by the addition of ATP in the absence of LC and is still able to fold, therefore, the role of BiP in antibody folding and assembly is to retain the HC in the ER until it associates with LC to ensure neither chain is secreted without its partner (Lee et al., 1999). The association of LC with the HC at the Fab region initiates the folding of the CH1 domain which is specifically carried out by the CL domain of the LC where it has been shown to act as a template and introduce structure into the CH1 domain of the HC *in vitro* (Feige, Groscurth et al., 2009). With the aid of other ER chaperones such as PPlases and PDI family members to catalyse the interchain disulphide between the CH1 and CL domains the complete folded IgG can then be secreted from the ER.

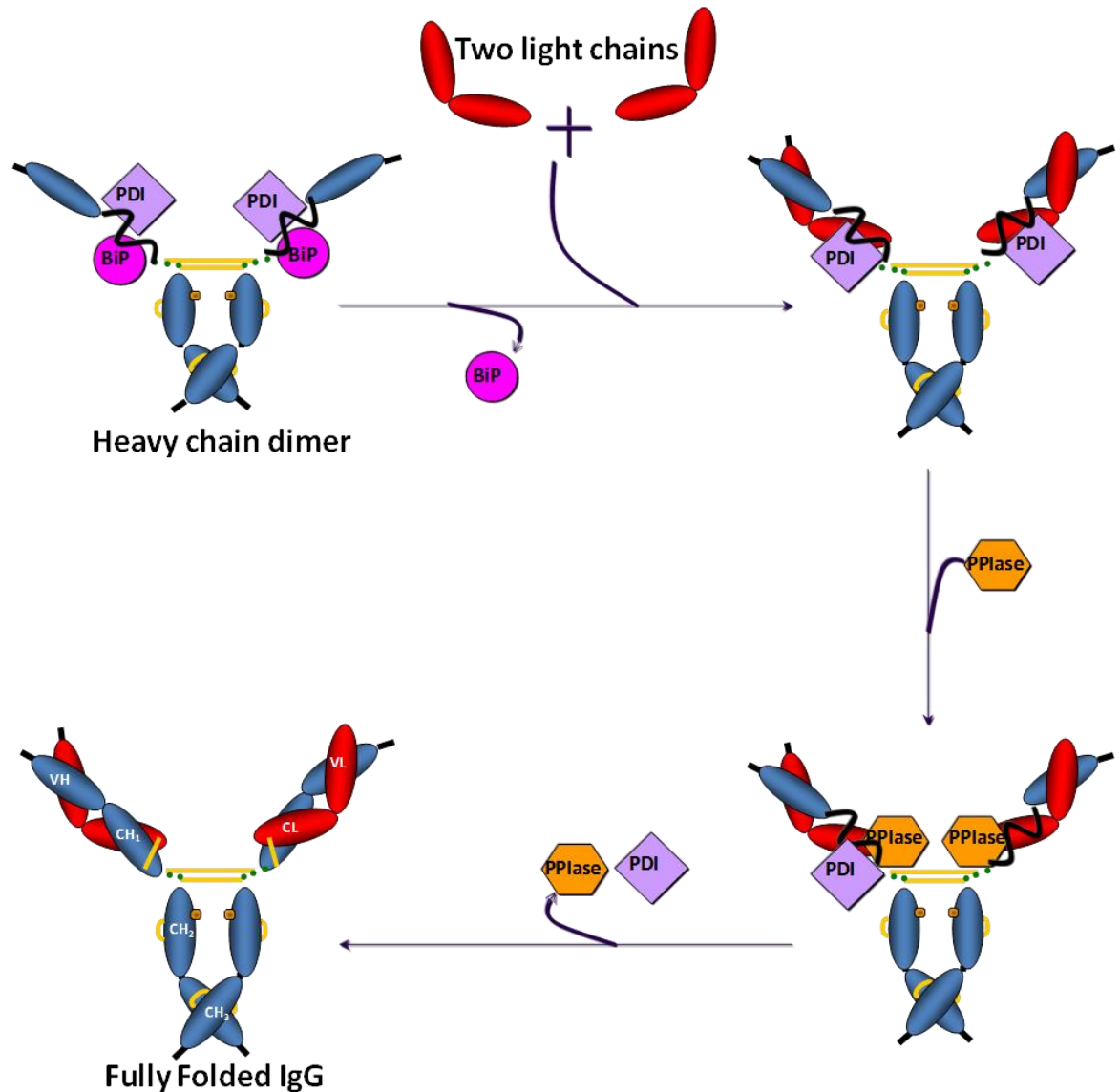


Figure 1.8: Antibody folding and assembly in the endoplasmic reticulum.

The LC (red) and HC (blue) of IgG are co-translated into the ER and folding of the immunoglobulin domains is initiated. The CH1 domain within the HC remains in a disordered, unfolded state and is sequestered in the ER via the action of the chaperone BiP (pink). Upon interactions of the variable domains, BiP is displaced by the CL domain of the LC and initiates folding of the CH1 domain. A PPIase (orange) then acts on a proline residue within the CH1 domain to isomerise it from a *trans* to a *cis* conformation. The action of PDI (purple) introduces the interchain disulphide bond (yellow) between the CL and CH1 domains forming a completed IgG folded structure which can then exit the ER. The CH2 domains of the HC are also glycosylated (orange) via an asparagine residue prior to exit from the ER.

The folding of the V regions is difficult to characterise because, although they form the same immunoglobulin fold, the large sequence variations between them can alter the folding of any specific V domain. For example, the VL domain of MAK33 (a murine IgG1) has been shown to follow two potential folding pathways; one demonstrates a secondary structure leading to its native state following the isomerisation of proline which is the standard folding pathway of immunoglobulin domains, the second pathway produces a more native intermediate prior to the protein folding to its native state, this intermediate has been implicated in the formation of amyloids (Simpson, Herold et al., 2009). This folding pathway was also the case for the VL domain of an antibody which binds ferritin, where upon refolding two conformations of the VL domain were formed; one being more functional than the other (Tsybovsky, Shubenok et al., 2007). Alternatively, the VH domain of the anti-ferritin antibody demonstrated a single folding pathway with a partially structured, yet functional native state (Martsev, Dubnovitsky et al., 2002). It has been demonstrated that the stability of the VH domain can be increased by mutating some of the sequence to more resemble that of a camel VH domain (description of camelids see section 1.3.4) with the addition of a second intrachain disulphide bond between the two β sheets in addition to other stabilising point mutations (Davies & Riechmann, 1996). The differences in stability and folding of variable domains complement each other when the two are expressed together. Investigations into the framework sequences (sequences of the variable domain between the CDRs) have highlighted similarities which contribute to the association of the VL domains with the VH domain. For example, the alignments of VL and VH interactions have shown ionic interactions between arginine and aspartic acid plus paired glutamine residues which aid in strengthening the domain interactions of the variable region, furthermore a conservation of aromatic residues has been implicated in the assembly of each immunoglobulin domain (Campion, 2016). The conserved aromatic residues are important for the association of the VL and VH domains as they protrude from the β sheets of the opposite domains to form a “packing” structure between the two domains providing a sheet to sheet assembly that is different for other IgG domain dimers (Chothia, Novotný et al., 1985). The hydrophobic interactions of the VL and VH domains also contributes to a strong association between the two domains as it was found that if the first portion of the J segment (residue 96) of the LC was hydrophobic as opposed to charged the association with the VH domain was stronger, the size and hydrophobicity of the residue also correlated with association (Hamel, Isenman et al., 1984). It has also been shown that the H bonds between the VH and VL domains are important in associations between variable domains, but not essential. For example, an investigation into the highly conserved H bond between a pair of glutamines (one on the VH domain and one on the VL domain), demonstrated that this H bond is not essential but association and stability can be improved via site directed mutagenesis to alter

the interaction to an ionic one rather than a H bond; this investigation was carried out on a single chain Fv (scFv), however, and may not be true for variable domains as part of a complete IgG protein (Tan, Sandmaier et al., 1998). The stabilising interaction and strong affinity of the VH and VL domains for each other enables association of the CL and CH1 domains for complete antibody folding. Therefore, the interactions of the variable domains govern the initial folding and contribution of the complete globular structure of IgG.

1.3.4 LC and HC dimer secretion

The assembly of the immunoglobulin chains which is tightly regulated by BiP and the association of the heavy and light chains via their variable domains contributes to the production of a fully folded IgG protein with two identical HCs and two identical LCs. However, there are a few exceptions to this folding pathway which results in the secretion of HC and LC separately rather than part of a complete immunoglobulin protein. The secretion of LC dimers was found to be dependent on the VL domains associating with each other and forming covalent interactions to be secreted; conversely VL domains that did not interact with one another were instead retained in the ER via BiP and eventually degraded (Leitzgen, Knittler et al., 1997). The occurrence of HC dimer secretion in nature exists for both camels and sharks. Sharks have adopted a different immune system to humans and instead of antibodies they possess immunoglobulin/T cell receptor like proteins consisting of a single variable and five constant domains which are secreted into serum in the absence of a partner LC-like protein (Greenberg, Avila et al., 1995). Camels have HCs which lack the CH1 domain, and instead have extended hinge regions; this leads to the secretion of HC dimers as they do not possess a disordered domain for BiP to bind to (Hamers-Casterman, Atarhouch et al., 1993). The VH domain of the camel HC is also very stable which can contribute to lack of association with its partner VL domain (Muyldermans, Atarhouch et al., 1994).

Investigations into HC dimer expression and secretion in humans for both natural and synthetic constructs have also been investigated. There have been reported cases of mutant HCs of IgM being expressed on the cell surface of pre-B cells without its partner surrogate LC (Minegishi & Conley, 2001). This was also reported for a different mutant IgM HC which was able to fold and be expressed on the cell surface in the absence of LC, which questions the previously outlined mechanism for B cell development as both LC and HC are required for the development of B cells (Shaffer & Schlissel, 1997). Synthetic production of antibodies using CHO cells also demonstrated the secretion of HC dimers when the ratio of LC and HC expression was altered, however, this was attributed to aggregation which induced stress of the cell with the ER going over capacity of the

protein within it and concluded the HC dimers secreted were incorrectly folded (Ho, Wang et al., 2015).

1.4 Antibodies and novel antibody formats used as therapeutics

1.4.1 Humanisation of antibodies

The function of antibodies in the immune system has been manipulated to generate synthetic proteins which can be utilised as therapeutics. The high specificity of the variable domains of immunoglobulins can be utilised to target specific epitopes which are desired for therapeutic use, in combination with the immune response. This approach comes with several advantages over using small molecules as drugs as the side effects will be greatly reduced in comparison to the non-specific binding effects observed from drugs which are already widely used. One area where therapeutic antibodies are already extensively used is the treatment of several cancers and immune deficiencies (reviewed in (Scott, Wolchok et al., 2012). Cancers are difficult to treat as they are not recognised as a foreign attack on the body as the cells which form tumours are from the host, therefore specifically attacking these cells over healthy non-cancerous cells is difficult to achieve. However, there are subtle chemical markers which are presented specifically on cancer cells which create potential antibody-binding sites generating a specific site of attack which does not affect healthy tissue. The engineering potential allows for the design of antibodies to be used as agonists, antagonists, or a delivery system for a drug conjugated to the protein or using the immune response initiated via antibodies. The generation of these antibodies from animals for human use brings with it the issues of producing an immune response against the antibody if it is recognised by the host as foreign rather than attacking the target antigen. In order to overcome this issue potential therapeutic antibodies undergo a humanisation process before being introduced as a therapeutic for human use.

The humanisation process begins with an animal (usually mouse) inoculated with the antigen of choice to enable production of antibodies with tight binding and high specificity. The animal's blood is screened for production of high titres of antibodies in the serum via ELISA method (Ward & Antibodies, 1999). The splenic cells of the animal are then extracted and fused with a myeloma cell line to generate a hybridoma cell which can be screened for the antibody (de StGroth & Scheidegger, 1980, Köhler & Milstein, 1975). Identification of the antibody involves testing clonal populations of the hybridoma cell line which have successfully fused via purification of the antibody produced in the serum and testing its affinity with the antigen used during inoculation (Ozato, Mayer et al., 1980). Once a cell line has been generated which successfully expresses the monoclonal antibody, the sequence of each variable domain can be obtained via PCR providing the data for the CDRs which have been developed during antibody evolution in the B cells of the

animal. Therefore, using DNA recombinant technology, the development of the humanisation of antibodies arose to avoid an unwanted immune response.

The initial development of humanised antibodies was to create chimeric proteins which consist of the animal variable domain and antigen binding site fused onto human constant domains of a desired antibody isotype (Morrison, Johnson et al., 1984). By manufacturing a recombinant DNA sequence of the variable region of an animal which leads onto the DNA sequence of the human antibody constant domains, a synthetic antibody product can be generated and purified for use as a therapeutic (Boulianne, Hozumi et al., 1984). However, the variable domain of the chimeric antibody still contains foreign non-human framework sequences between the CDRs; these sequences still have the potential to initiate an immune response against the therapeutic regardless of the human constant domains. Therefore, further engineering of the variable domains is required. There are two different techniques available to humanise the variable domain; one involves grafting the CDR regions onto a human framework backbone, and the other involves making point mutations within the framework sequence so that it reads as human, these point mutations are also important in the overall structure and stability of the variable domain (Ewert, Huber et al., 2003, Jones, Dear et al., 1986). However, combinations of both are usually required as in some cases the framework sequence is also involved in antigen binding to a small extent in contribution with the CDRs (Ewert, Honegger et al., 2004). This process can be manipulated to design a variable domain with a specific level of binding affinity (high for therapeutics, low for purification), for example, the antigen binding of an antibody for lysozyme was improved via mutations of the variable domain at specific sites using a modelling and site directed mutagenesis approach (Roberts, Cheetham et al., 1987). The process of CDR grafting uses antibody framework sequences that are similar between the animal and human in order to provide a chemically favourable environment, however, the opposite has also been investigated where instead of looking for similar framework sequences, the CDR sequences were compared between human and animal as it would already provide a chemically suitable environment for the chosen CDRs (Hwang, Almagro et al., 2005).

The development of humanisation is an on-going process, as producing a fully human antibody is unachievable, and CDR grafting still results in immunogenicity as the immune system recognises small peptides as foreign which in the case of humanised antibodies could be the CDRs themselves or the junctions where the CDRs are fused with the human framework (Clark, 2000). Therefore, an alternative process known as human string content (HSC) has been developed. HSC involves quantifying the small peptide sequences within the variable domain with known human peptide sequences to determine the potential level of immunogenicity which could occur, providing a more accurate level of engineering when humanising the variable sequences (Lazar,

Desjarlais et al., 2007). After the humanisation process is complete, binding assays to the antigen target are carried out to confirm any changes made do not inhibit the function of the therapeutic in binding to its antigen. This process generates other possibilities of fusion of variable domains to other proteins such as enzymes depending on the therapeutic action required (Williams & Neuberger, 1986) (see section 1.4.3 for further discussion).

1.4.2 Generation of therapeutics using CHO cells

Once the sequence of a humanised therapeutic antibody has been determined, a production system which ensures correct assembly of the protein product must be chosen carefully. A mammalian system is required as it contains the essential chaperones and folding environment necessary for correct protein formation and addition of post-translational modifications. Chinese hamster ovary (CHO) cells are the most frequently used mammalian cell line for the mass production of therapeutic proteins. CHO cells were initially developed in the 1950s as a laboratory cell line which today are utilised for industrial scale cultures which produce therapeutic proteins as the products have shown to be compatible with human use and the cells are able to produce high levels of protein (Jayapal, Wlaschin et al., 2007). For this reason, the engineering of CHO cells is still in development today to improve on the yield and quality of protein produced.

When using CHO cells for protein production, there are several potential engineering avenues to consider. For example, over-expressing a protein in a mammalian system could result in aggregation; in the case of antibody production the ratio of LC and HC, cell density, folding chaperone expression and efficient protein folding in the ER are all factors to consider to prevent protein aggregation and cell death (Ishii, Murakami et al., 2014). The ER organelle and its chaperone content have been artificially engineered to incorporate positive attributes that the cells displays when under stress and the UPR is activated (Pybus, Dean et al., 2014). There is a fine balance between recovery of protein folding and inducing cell death upon activation of the UPR, as this process has been well studied, specific proteins have been identified and overexpressed which contribute to a higher level of protein folding efficiency. One example of a component of the UPR which has been overexpressed to produce higher levels of antibody is the transcription factor XBP1 which results in an increase in the size of the ER and overexpression of chaperones and foldases in the ER which has been shown to increase protein productivity and decrease aggregation (Ku, Ng et al., 2008). When XBP1 is co-overexpressed with the ER chaperone Ero1 antibody production is increased further demonstrating the importance of the folding environment which is engineered in CHO cells when producing monoclonal antibodies (Cain, Peters et al., 2013).

1.4.3 Antibodies and novel antibody formats developed for therapeutic use

Investigations into antibody structure and function has developed into the generation of novel antibody formats and conjugates with the exploitation of the positive attributes of antibodies with tailored functions for different therapeutic uses. There are currently over 60 different antibody formats in development which is a relatively new area of antibody engineering as the success of therapeutic monoclonal antibodies and conjugates has enabled further research and development of this area of pharmaceuticals (Spiess, Zhai et al., 2015).

One example where antibody conjugates are already widely used is for the treatment of several cancers. Conjugation of a cytotoxic drug to an antibody that recognises a specific epitope on a tumour cell has been met with great success, for example Herceptin (trastuzumab) which is used in the treatment of metastatic breast cancer. Herceptin is a humanised monoclonal antibody which binds to ErbB2 an epidermal growth factor receptor which is involved in the signalling pathway of metastatic breast cancer, binding of Herceptin inhibits tumour growth and has had huge success in treating the disease (Baselga, Tripathy et al., 1996). With this success, however, some patients unfortunately still progressed with the disease, therefore trastuzumab was further developed by conjugating a microtubule-depolymerizing agent (a cytotoxin which inhibits mitosis) via disulphide and thioether linkers to generate trastuzumab-maytansinoid which was seen to be a more effective treatment than the monoclonal antibody alone (Lewis Phillips, Li et al., 2008). Another example is the conjugation of ricin via a disulphide linker to a murine monoclonal antibody J-5 which binds to CALLA an antigen present on acute lymphoblastic leukaemia cells. The conjugated J-5 antibody was tested initially using a rabbit-reticulocyte lysate system and later *in vivo*, the results showed a 50% decrease in proliferation and a 70% increase in efficacy when compared to the antibody treatment alone (Raso, Ritz et al., 1982). Therefore, the addition of a cytotoxic conjugate greatly increases the effectiveness of the antibody while maintaining its specificity. Another cytotoxin conjugate used in cancer therapy is radiation, whereby the antibody is labelled with ionizing radiation which is targeted to tumour cells and internalised enabling specific action of radiation while avoiding emitting the dosage to healthy tissue (Milenic, Brady et al., 2004). The most commonly used radionuclide in the treatment of radioimmunotherapy is Iodine-131, however the half-life and effect of this radiolabel is compromised due to swift deiodination by enzymes due to similarities in structure with thyroid hormones (Zalutsky & Narula, 1987). However, the introduction of a linker (IBPA) containing the iodine radiolabel conjugated to cetuximab (a therapeutic antibody) showed a higher level of

internalisation and a lower level of deiodination, demonstrating the effectiveness of linkers in therapeutic action of monoclonal antibodies (Kim, Kim et al., 2016).

Alternatively, another treatment for cancers is the recruitment of T cells to target and kill tumour cells. This area of development of novel antibody formats is the generation of bi-specific antibodies, whereby the Fab region consists of two different variable domains capable of recognising two different targets (Nisonoff & Rivers, 1961). For example, blinatumomab which is a bi-specific antibody which targets the tumour specific antigens CD19 and CD3 on acute lymphoblastic leukemia B cells, in conjunction with recruitment of cytotoxic T cells as its method of action (May & Glode, 2016). However, the generation of bi-specific heterodimer antibody formats has proved difficult as when expressed, the HCs preferentially associate and dimerise to form homodimers. Therefore, there have been several engineering developments which have shifted the equilibrium to the formation of heterodimers. One engineering opportunity was to manipulate the interaction between the CH3 domains in a 'lock and key' type mechanism to alter which HCs are able to dimerise (Ridgway, Presta et al., 1996). Alternatively, the Fab arm exchange which is demonstrated by IgG4 can be used to generate a bispecific antibody format. Fab arm exchange, where the two IgG chains dissociate under reducing conditions, occurs due to the contributions of the core hinge and CH3 domain of IgG4. The IgG4 sequences conferring Fab exchange were engineered into IgG1 and the generation of a bispecific antibody using Fab arm exchange was achieved (Strop, Ho et al., 2012). Another engineering route for the generation of bispecific antibodies is via manipulating the interactions of the variable domains to construct specific VH-VL pairings by re-designing the interface so that only one is paired with another. One approach for generating specific VH-VL pairings involves a collaboration of crystallographic and computational modelling techniques to design an orthogonal interface to make direct pairings to produce a stable and functional bi-specific therapeutic antibody (Lewis, Wu et al., 2014). Other approaches have used site directed mutagenesis to alter the design of the variable domains to produce exact VH-VL pairings which are wanted (Igawa, Tsunoda et al., 2010).

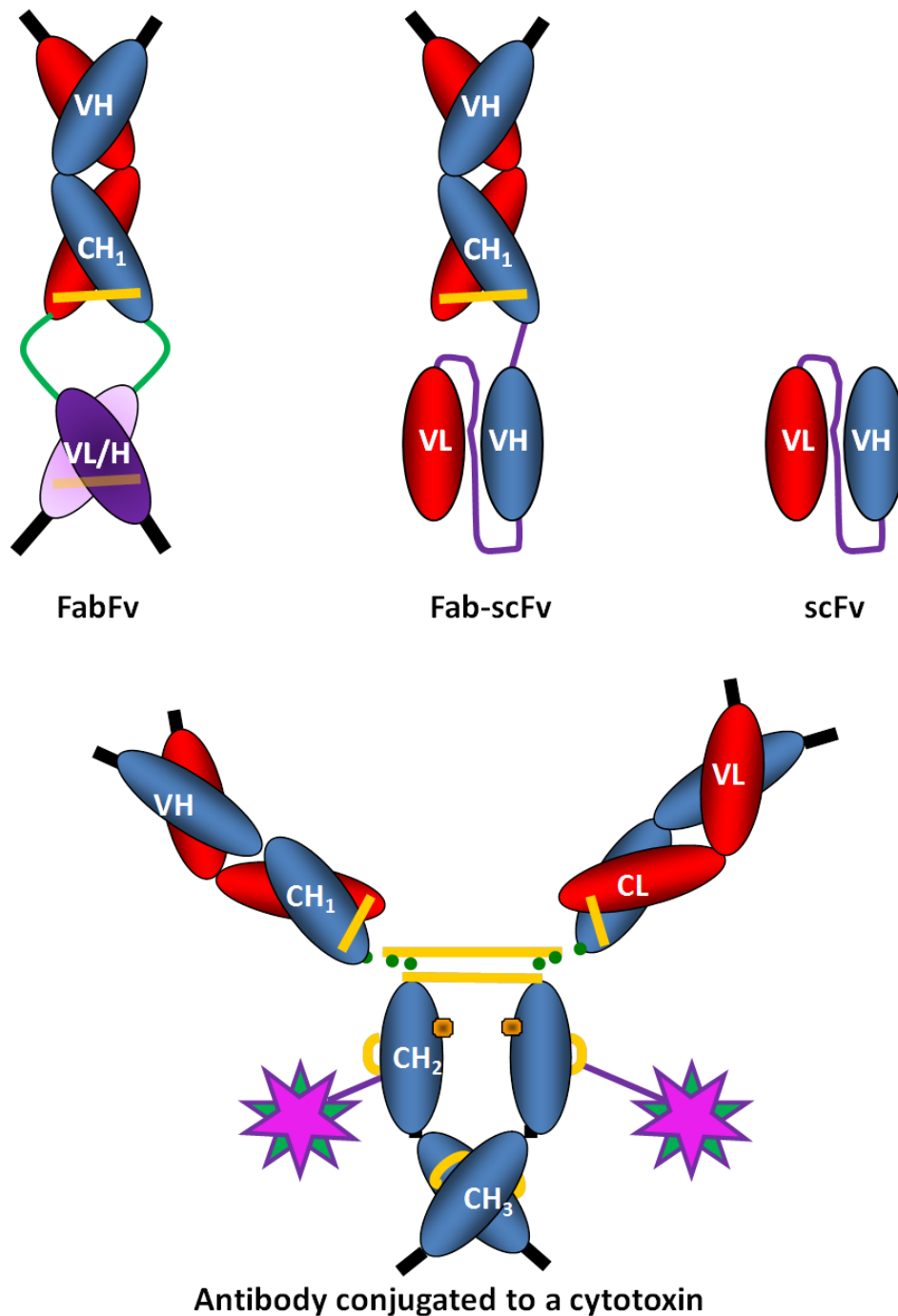


Figure 1.9: Schematic representation of a selection of various antibody formats utilised as therapeutics.

This figure depicts the structures of four antibody formats which are used as therapeutics. The FabFv consists of a HC (blue) and a LC (red) with a linker (green) to an additional variable domain (purple). The FabFv has an additional engineered disulphide within the second variable domain (purple). The Fab-scFv consists of a Fab fragment with a linker (purple) to a scFv which comprises of a VL and VH domain stabilised via a linker (purple). The scFv also exists as a therapeutic alone. All antibody formats are derived from the original antibody structure utilising different aspects of the domains for tailored therapeutic action. Depicted is also an example of a therapeutic antibody with a conjugated cytotoxin which administers a drug upon action of the therapeutic antibody.

There is a large repertoire of potential antibody formats which can be generated depending on the function of the therapeutic required, some examples discussed in this section are depicted in Figure 1.9. Examples include the addition or substitution of constant domains, single chain Fvs (scFvs) which are only comprised of the variable domains, Fabs, diabodies (two Fabs/two Fvs/two IgGs etc), tribodies (three antibody targets consisting of three Fvs), fusion antibodies (diabodies and tribodies) and many more (Spiess et al., 2015). Advantages for tailoring novel antibody formats are that they are more economical as they are easily produced in *E.coli* expression systems as well as more functional as a pharmaceutical to exploit only the functions required (Carter, 2006). However, some disadvantages are that the serum half-life is reduced when the Fc fragment is removed and the stability is also reduced in comparison to complete IgG. One example of the difference in functionality of different antibody formats is the investigation and comparison of the efficacy of a scFv, Fab and full antibody in the treatment of rheumatoid arthritis (RA) with the antigen target IL-1 β which is a cytokine involved in the inflammatory pathway associated with RA. It was found the Fab and full antibody were equally effective in treating the mice suffering from collagen-induced arthritis, and together more effective than the scFv, however, they deemed the Fab format as the therapeutic of choice as it is more economical to produce than the full antibody (Qi, Ye et al., 2014). The scFv has been found to be less stable than its counterpart Fab fragment, therefore, to utilise both qualities a novel bispecific antibody format was generated which fuses a scFv to either the HC or LC of the Fab fragment producing a Fab-scFv which recognises two antigen targets. The Fab-scFv is more stable than scFv alone and does not contain the CH2 and CH3 domains of the full antibody yet retains the binding efficacy of both parent antibodies in a cooperative manner (Lu, Jimenez et al., 2002). The overall stabilities of scFvs in comparison to Fab fragments is considerably lower, therefore, to overcome this instability, peptide linkers were introduced. The replacement of constant domains with a helix peptide linker has resulted in greater stability more comparable with the Fab counterpart (Arndt, Müller et al., 2001). The use of scFvs has an advantage over Fab fragments as they are smaller molecules and have found to penetrate the cell more readily than the larger counterpart therapeutic antibodies (Rothlisberger, Honegger et al., 2005). The engineering potential for several antigen targets, presence of the Fc fragment to activate ADCC and the complement system, composition of different chains and therapeutic potentials using the information already known regarding antibodies and their domains gives rise to a huge potential of possible therapeutics.

1.4.4 FabFv structure and function

A novel antibody format developed by the industrial company UCB Pharma is the FabFv. FabFvs consist of a single LC and a single HC. Each chain comprises of a variable and constant domain (Fab) joined to a single variable domain (Fv) via a 15 amino acid long glycine and serine polypeptide linker (Figure 1.9). The FabFv is a bispecific novel antibody format as it contains two variable domains. The overall structure contains two interchain disulphide bonds; one between the CL and CH1 domains which is already present in the quaternary structure of IgG, the second is an engineered interchain disulphide bond between two cysteines within the VL and VH domains of the second variable domain to enhance the overall stability of the FabFv structure. The FabFv was developed as an alternative bispecific therapeutic antibody which avoids activating the complement system as it does not contain the Fc fragment of the antibody. However, when the FabFv was expressed in CHO cells, the protein formed multimer species of FabFvs (UCB Pharma, unpublished). As the FabFvs are a novel protein, the folding and assembly of these antibody formats are still unknown.

1.5 Summary and project aims

The development, engineering and manufacture of therapeutic antibodies are essential areas of research to generate a novel type of pharmaceutical with the goal of increasing the efficiency of how the antibodies are produced. In order to increase the yield of antibodies and novel antibody formats using a CHO cell system, the folding and assembly of these proteins is a required area of investigation in order to understand how antibodies fold and what products are secreted.

Therefore, the project aims of this thesis were to investigate the folding and secretion of both monoclonal antibodies and FabFvs with the following hypotheses.

Hypothesis one: *Engineering of antibodies for therapeutic use alters their assembly and secretion.*

The first aim of this thesis was to investigate the folding, assembly and secretion of various antibodies to identify any areas of improvement which would ultimately increase the overall yield of antibody production. By focusing on the expression, folding and assembly of different immunoglobulins, this would enable the accomplishment of the second aim in gaining a greater understanding of how antibodies fold. All immunoglobulins investigated have undergone varying degrees of engineering and humanisation. The investigations in this thesis have determined any consequences to expression, folding or assembly that have been hypothesised to be as a result of engineering of the variable domain. Specifically, the results of cleavage of a HC expressed in CHO cells and the secretion of HC dimers have been investigated to note any consequences engineering of the VH domain have had on antibody expression and secretion.

Hypothesis two: *Multimer formation of the novel antibody format FabFv is a consequence of antibody engineering.*

Another area of examination of this thesis has focused on the folding and assembly of the novel antibody format FabFv. As very little is known about the expression and assembly of this novel therapeutic, initial investigations of FabFvs has aimed to increase the understanding of domain interactions of IgG during folding and assembly to again improve the yield of various therapeutics which are produced.

The knowledge gained from these investigations aim to improve on the engineering process of antibodies and novel antibody formats intended for therapeutic use.

Chapter 2 : Materials and Methods

2.1 Cell Culture

2.1.1 Cell maintenance

Chinese hamster ovary cells (CHO L761H) were grown in Dulbecco's Modified Eagle Medium (DMEM) (addition of 10% (v/v) FBS + 2 mM glutamine + 50000 units penicillin + 50 mg streptomycin + non-essential amino acids (10 μ M for each amino acid)) at 37°C under 5% CO₂. Cell line was kept in 75 cm² filter capped flasks and split into 6, 10 or 15 cm dishes and grown to approximately 70% confluency before use. Cells were washed in 1 X PBS (Gibco) prior to passage and trypsinised with 0.05% Trypsin-EDTA (Gibco) for 1 min at 37°C.

Chinese hamster ovary suspension cells (CHO-S) (Life Technologies) were grown in CD CHO chemically defined medium (addition of 4 mM glutamine) at 37°C under 5% CO₂. This cell line was grown in 125 ml plain bottom vented shaker flasks and seeded at approximately 200,000 cells/ml for routine passage and at approximately 500,000 cells/ml for a transfection. Cells were shaken at 120 rpm.

All cells were passaged every 3-4 days. The passage for each experiment was carried out using the same number of cells and incubated for the same length of time prior to transfection and further experimentation. Cells were not passaged past a passage of 20 as transfection efficiency dropped.

2.1.2 Transient transfection of DNA constructs

All constructs used in this thesis are outlined in table 1. The protein complexes produced and relative molecular weights from the constructs are outlined in appendix 1.

Two methods of chemical based transfection were used; initially polyethylenimine (PEI) (Durocher, Perret et al., 2002) and alternatively NovaCHOice (Novagen) (Slanina, Mündlein et al., 2014) as an substitute when transfection efficiency was poor.

PEI: CHO cells were passaged and transfected the following day. Alternative quantities of DNA was used according to cell number/density (4 μ g 6 cm dish/ 8 μ g 10 cm dish/ 16 μ g 15 cm dish). The ratio of DNA:PEI for all transfections was 1:2.5. DNA was initially added to serum-free media and incubated for 5 min. PEI was then added to the mixture, vortexed and incubated for a further 10 min. The mix of DNA and PEI was then added to cells in antibiotic-free media. The media was replaced with complete media (serum plus antibiotic) after 24 h. The transfected cells were incubated for 48 h post transfection prior to any experimental protocols.

NovaCHOice: CHOS cells were passaged and transfected the same day, CHO L761H were transfected the following day as it is an adherent cell line. A cell density of 500,000 cells/ml was

used with 20 µg DNA and 20 µL NovaCHOice. The DNA and NovaCHOice were added simultaneously in serum-free media and incubated for 10 min. The mix was then added to the cells and incubated for 48 h prior to experimental procedure.

2.2 DNA Methods

2.2.1 Site-directed mutagenesis

Appropriate primers were designed from original sequences for all mutants made (see appendix 2). All IgG4 antibodies contain a S228P mutation to stabilise the hinge region of the antibody unless stated otherwise. Mutations introduced were as follows: C131S in the A33 (VH) isolated Fab fragment, P151A in the A33 (VH) full length HC and P228S in the A33 (VH) full length HC. All mutations followed the same site-directed mutagenesis approach under the following conditions using the individual primers listed in table 4.

PCR was performed for 25 cycles with the designed primers at a T_m of 60°C. The composition of each PCR reaction was 50 ng of template DNA in 1 X Accuzyme buffer (60 mM TrisHCl, 6 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, pH 8.3) (Bioline), 2 mM MgCl₂ (Bioline), 10 µM of each primer (forward and reverse), 10 mM dNTP mix, 2% (v/v) DMSO, 2.5 units Accuzyme (Bioline) plus H₂O to give a final volume of 50 µL. The cycling parameters of the reaction were as follows; 95°C for 1 min, 60°C for 1 min, 72°C for 18 min. After completion 10 units *Dpn 1* (NEB) enzyme was added to digest the parental DNA which had not been mutated. The sample was then transformed into competent XL1 BLUE *E.coli* cells and positive colonies were selected on antibiotic (either kanamycin or ampicillin depending on the vector requirements at 50 µg/ml) agar plates. Mini and midi preps were carried out by growing a picked colony overnight in 2 ml (mini) or 100 ml (midi) of Luria broth (see section 2.2.3).

2.2.2 Sub cloning of DNA fragments into vectors by PCR or restriction digests

Plasmid maps of the vectors pMhg4PFL, pc DNA 3.1 and pET28a with the restriction sites used are outlined in Figure 2.1.

To generate the A33 IgG1 hinge mutant the following vectors were used for subcloning; the first was IgG4 A33 HC (UCB) (vector) and the second was IgG4 (M33) (insert) which was made by Shirley Peters (UCB). Both plasmid maps had the same restriction enzymes at the two sites which contained the CH₁ domain and hinge region of the antibody. The vector and insert were incubated separately at 37°C with 20 units EcoRI and 20 units XhoI for 2 h consecutively, denaturing the first enzyme at 70°C for 10 min before adding the second. After the digest, 5 µL of

6 X DNA loading dye (10 mM Tris-HCl (pH 7.6) 0.03% (w/v) bromophenol blue, 0.03% (v/v) xylene cyanol FF, 60% (v/v) glycerol 60 mM EDTA) (Thermo Scientific) was added and the sample was loaded onto a 1% (w/v) agarose gel (1g agarose in 100 ml 1 X TAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA) plus 10 µL SYBR Safe DNA gel stain (1:10000) (Invitrogen) for gel electrophoresis. Electrophoresis was performed for 60 min at 80V in 1 X TAE buffer. A 1 KB marker (5 µg/µL) (Invitrogen) was used to size the DNA fragments which were visualised under UV guidance. The appropriate insert and vector DNA fragments were excised from the gel and weighed. The DNA fragments were purified from the gel using a Qiagen Quick Gel Extraction kit as per the manufacturer's instructions. Once purified, the DNA concentration was determined by measuring the OD at 260 nm using a Nanodrop (Thermo Scientific) or Spectrostar nano (BMG Labtech). Any protein contaminants of the sample were determined by an OD 260:280 ratio that provided a value between 1.8 and 2.0. Ligation reactions contained 100 ng of vector with either 3 or 5 times the insert using the following calculation:

$$\frac{ng\ of\ vector \times kb\ size\ of\ insert}{kb\ size\ of\ vector} \times \frac{molar\ ratio\ of\ insert}{ng\ of\ vector} = ng\ of\ insert$$

The ligation reactions consisted of insert + vector in ligation buffer (250 mM Tris-HCl (pH 7.6), 50 mM MgCl₂, 5 mM ATP, 5 mM DTT, 25% (w/v) polyethylene glycol-8000) (Invitrogen) with 1 unit T4 DNA ligase (Invitrogen) with a final volume of 10 µL were performed at 16°C and incubated overnight. The following day each ligation reaction was transformed into XL1 blue competent cells and plated out onto antibiotic agar. Mini preps were undertaken for colonies which were present and confirmation of a successful ligation was verified by restriction enzyme digest. Final confirmation of positive sub cloning was determined via sequencing of the DNA by GATC biotech.

The vector containing the sequence of A33 (VH) with the full length IgG4 cDNA was used to swap the coding region of the variable domain for variable regions HC1, HC2, HC3, HC4 and HC5 using the restriction sites Hind III and XhoI. The same sub cloning protocol using restriction enzyme digestion was used as for the IgG1 hinge construct.

To generate constructs for the expression of individual in *E.coli* immunoglobulin domains (A33 VH domain, CH1 domain, CL domain) the relevant sequence was first synthesised commercially (Genescript). All constructs were codon optimized for expression in *E.coli*. To subclone the DNA sequence of each domain into an appropriate expression vector (pet28) the coding regions were first amplified using the primers which are listed in Table 4 which also appended the restriction enzyme sites for NdeI and Hind III at the 5' and 3' ends respectively. PCR was performed under the same conditions described in section 2.2.1. The PCR product underwent gel electrophoresis on a 1% (w/v) agarose gel in 1 X TAE buffer. The PCR product was extracted from the gel using the Qiagen Quick Gel Extraction protocol and subsequently digested with 20 units of the

restriction enzymes Nde1 and Hind III for 1 h consecutively, denaturing the first enzyme before incubating with the second. The pET28a vector was digested with restriction enzymes Nde1 and Hind III as described previously. The insert and vector digested DNA was extracted, purified, ligated, prepared and sequenced as described above.

Generation of the truncated HC A33 Fab construct was carried out using the primers A33 Fab (F and R) (appendix 2) to amplify the cDNA sequence of the VH and CH1 domains from the A33 IgG4 full length HC sequence in the pMhg4PFL vector. These primers kept the original Hind III restriction site at the 5' end and added the restriction site Not1 onto the 3' end for sub cloning into an empty pcDNA 3.1 vector. PCR and restriction digests were carried out as previously described.

Antibiotics used for colony selection were kanamycin (50 µg/ml) for expression vectors pMhg4PFL and pet28, and ampicillin (50 µg/ml) for the expression vector pc DNA 3.1.

Table 1: Table of constructs used and generated in this thesis.

All mutations numbering starts from V sequence of A33 IgG4 except * which is IgG1.

Construct	Antibody Isotope	Species of V domain	Mutation	DNA sequence	Vector	Tag	Source
A33 LC	Kappa	Mouse-human chimeric	WT	Mammalian	pMhg4PFL	-	UCB
A33 HC	IgG1	Mouse – human chimeric	WT	Mammalian	pMhg4PFL	-	UCB
A33 HC	IgG4	Mouse – human chimeric	S225P	Mammalian	pMhg4PFL	-	UCB
A33 HC	IgG4	Mouse – human chimeric	WT -P225S	Mammalian	pMhg4PFL	-	UCB
A33 HC IgG1 hinge	IgG4	Mouse – human chimeric	C127S S227P Y229S G230C P237D P238delinsKTHT *	Mammalian	pMhg4PFL	-	UCB
A33 HC	IgG4	Mouse – human chimeric	P151A S225P	Mammalian	pMhg4PFL	-	UCB
A33 HC	IgG4	Humanised	S225P	Mammalian	pMhg4PFL	-	DNA 2.0 (VH)
A33 HC Fab	IgG4	Mouse – human chimeric	WT	Mammalian	pc DNA 3.1	V5	UCB
A33 HC Fab	IgG4	Mouse – human chimeric	C131S	Mammalian	pc DNA 3.1	V5	UCB

A33 LC FabFv	Kappa	Mouse – human chimeric	WT	Mammalian	pMhg4PFL	-	UCB
A33 HC FabFv	IgG1	Mouse – human chimeric	WT	Mammalian	pMhg4PFL	-	UCB
HC1	IgG4	Human	S225P	Mammalian	pMhg4PFL	-	UCB
HC2	IgG4	Mouse – human chimeric	S225P	Mammalian	pMhg4PFL	-	UCB
HC3	IgG4	Humanised	S225P	Mammalian	pMhg4PFL	-	UCB
HC4	IgG4	Mouse – human chimeric	S225P	Mammalian	pMhg4PFL	-	UCB
HC5	IgG4	Humanised	S225P	Mammalian	pMhg4PFL	-	UCB
497 LC	Kappa	Humanised	WT	Mammalian	pMhg4PFL	-	UCB
497 HC gH8	IgG4	Humanised	S225P	Mammalian	pMhg4PFL	-	UCB
497 LC FabFv gH8	Kappa	Humanised	WT	Mammalian	pMhg4PFL	-	UCB
497 HC FabFv gH8	IgG1	Humanised	WT	Mammalian	pMhg4PFL	-	UCB
A33 VH domain	VH domain	Mouse	WT	<i>E.coli</i>	pET 28	His	Genescript
CH1 domain	IgG4	Human	WT	<i>E.coli</i>	pET 28	His	Genescript
CL domain	Kappa	Human	WT	<i>E.coli</i>	pET 28	His	Genescript

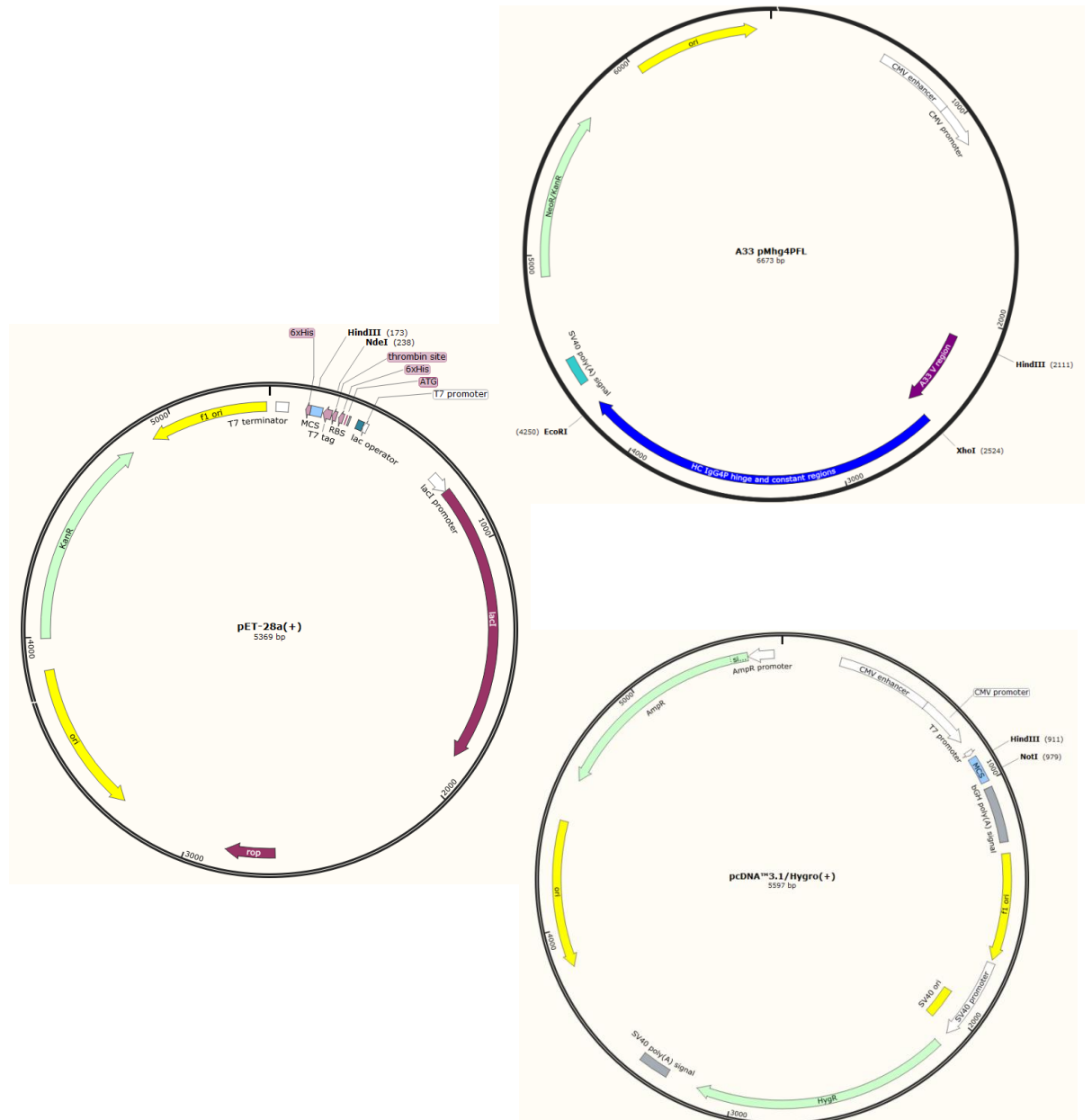


Figure 2.1: Plasmid maps of the vectors pMhg4PFL, pET-28a and pcDNA 3.1 used to generate each construct.

All full length HC antibody constructs used to generate HCs with an alternative VH domain or hinge region used the pMhg4PFL vector with the restriction sites Hind III and XhoI for the variable domain and XhoI and EcoRI for constant domains. This vector contains kanamycin resistance. The pET-28a vector was used to express the single immunoglobulin domain in *E.coli* using the restriction sites HindIII and NdeI. This vector also contains kanamycin resistance. The pcDNA 3.1 vector was used to generate the truncated A33 HC Fab fragment and used the restriction sites HindIII and Not1. This vector contains resistance to ampicillin.

2.2.3 Mini and Midi plasmid DNA purification from transformed *E. coli*

DNA was extracted from overnight cultures inoculated from a single colony using Qiagen Plasmid Midi Kit following the manufacturer's instructions. The mini prep was carried out by re-suspending pelleted cells in GTE buffer (25 mM TrisCl pH8, 50 mM glucose, 10 mM EDTA) with RNase A (10 mg/ml). After 2 min incubation at room temperature, 200 μ L of fresh NaOH/SDS (0.2 M NaOH, 1% (w/v) SDS) solution was added, incubated on ice for 5 min and then 150 μ L 5 M potassium acetate solution pH 4.8 was added. Cell debris and chromosomal DNA were pelleted and the supernatant was transferred to a solution of 450 μ L phenol:chloroform:isoamyl alcohol (25:24:1). The upper phase was then removed and transferred to 450 μ L chloroform. Finally, the aqueous phase was removed and added to 800 μ L 95% (v/v) ethanol, washed with 1 ml 70% (v/v) ethanol and air-dried. DNA was re-suspended in 30 μ L H₂O for sequencing.

2.3 Protein Methods

2.3.1 Western blot analysis

The following primary antibodies used during western blot and immunoisolation experiments during this thesis are summarised in Table 2.

Table 2: Table of antibodies with corresponding catalogue numbers used in this thesis.

Antibody	Dilution	Company	Catalogue number
Anti V5	1:10000	Invitrogen	R96025
Anti human κ LC	1:1000	Sigma-Aldrich	K4377
Anti Human IgG HC	1:1000	abcam	ab7500
Anti Calnexin	1:2000	Stressgen	ADI-SPA-600-D
Anti GAPDH	1:10000	Ambion	AM4300

Immunoisolation of the A33 Fab construct containing a V5 epitope tag was carried out using anti-V5 agarose affinity gel beads (Sigma) as an alternative to protein A Sepharose. Fluorescent secondary antibodies were used for the appropriate species of each primary antibody (Licor). A fluorescent protein A was used as an alternative secondary antibody or as a primary antibody when detecting IgG and HC products (Schellenberger, Weissleder et al., 2004) (Donna McGow, University of Glasgow).

Cells were transfected using PEI/NovaCHOice with various combinations of plasmid (as described in figure legends of results) and incubated for 48 h. Cells were incubated in phosphate buffered saline containing NEM (20 mM) for 10 mins on ice prior to lysis. After treatment and incubation, cell lysates were prepared using 120 μ L of lysis buffer (50 mM TrisHCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 1% (v/v) Triton X-100) with 20 mM NEM. Cell lysates were pelleted by centrifugation at 9500xg for 10 min to pellet the cell debris. The supernatant was extracted and pre-cleared using 10% Sepharose for 30 mins before affinity purification with protein A Sepharose beads overnight. Samples were washed with immunoisolation buffer (50 mM TrisHCl pH8, 1% Triton X-100, 150 mM NaCl, 2 mM EDTA, 0.5 mM PMSF) before adding 2X sample buffer (100 mM TrisHCl pH 6.8, 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol) to elute the protein from the protein A beads. Protein samples were separated by SDS-PAGE on a gradient (4-20%) polyacrylamide gel. Gel electrophoresis was carried out in 1 X running buffer (25 mM Tris, 20 mM Glycine, 3 mM SDS) for 2 h at 20 mA per gel. The gel was transferred onto nitrocellulose paper (Licor membrane) at 150 mA for 2 h in 1 X transfer buffer (25 mM Tris, 20 mM Glycine, 3.5 mM SDS, 20% (v/v) methanol). Blots were blocked for 1 h at room temperature in 5% (w/v) milk and incubated with a fluorescent protein A or primary antibody for 1 h. The blot was washed in 1 X TBST (20 mM Tris, 15 mM NaCl, 1% (v/v) Tween) and then incubated with the secondary antibody for 45 min (if fluorescent protein A was used as a primary antibody then no secondary was used). The blot was washed again in TBST prior to development. The development of the western blot was carried out using an Odyssey Licor Sa imaging system and fluorescence from the protein A and secondary antibodies was measured at 800 nm.

2.3.2 Pulse chase assay

Cells were transfected as described in section 2.1.2. After a 48 h incubation post transfection cells were starved in minus cysteine and minus methionine (-cys-met) DMEM medium for 30 min. The medium was replaced with fresh -cys-met with the addition of a radiolabeled methionine and cysteine mix (35 S) for 30 min at a concentration of 110 μ Ci/ml. Cells were then washed in PBS and 1 ml complete DMEM containing 0.5 mM cycloheximide was added to each dish. After treatment and incubation cell lysates were prepared as described in section 2.3.1 at various time points (indicated in Figure legends). Protein samples were immunoisolated using anti κ LC and anti γ CH₁ HC antibodies and incubated overnight with protein A Sepharose. The following day the beads were washed with IP buffer and eluted in sample buffer. Protein samples were separated by SDS-PAGE on a gradient (4-20%) polyacrylamide gel. Gel electrophoresis was carried out in 1 X running buffer for 2 h at 20 mA per gel. The gel was fixed using a solution of 10% (v/v) methanol and 10% (v/v) acetic acid for 20 min, and then dried on a DryGelSr for 1 h 10 min at 80°C. Prior to exposure, the pre-stained protein markers on the dried gel were labelled with a spot of 35 S met

diluted by 1 in 500,000 to make the marker visible when exposed. The dried gel was then exposed using a phosphorimager plate overnight and developed on a FLA-7000 phosphorimager.

The methionine content of the antibody chains investigated were equal for FabFvs (both contain 3) and more than doubled in comparison of LC to HC in the full length antibody (LC contains 2 and HC contains 5). These values were considered during quantification of radiolabelled proteins.

2.3.3 2D gel electrophoresis

Transfected cells were prepared in 15 cm dishes for radio-labelling. A pulse chase assay (as described in section 2.3.2) with a pulse time of 30 min using a radiolabelled methionine and cysteine mix (^{35}S) at a concentration of 110 $\mu\text{Ci/ml}$ and a chase time of 3 h was used. After 3 h the lysate and medium were collected separately. Once the samples had been prepared (cell lysis as described in section 2.3.1) and isolated with protein A Sepharose (plus anti κ LC and anti IgG antibodies to isolate all intermediates) they were loaded onto two separate gradient gels; one for the lysate samples and one for the medium samples. Gradient gels were run at 20 mA per gel for 2 h under non-reducing conditions as the first dimension. A sample lane was then cut (approximately 3mm wide) using a scalpel and incubated in 50 mM DTT to reduce the proteins in the gel. The other gel piece was fixed in 10% (v/v) acetic acid/ 10% (v/v) methanol. The gel lane that was incubated in 50 mM DTT for at least 10 min was then placed horizontally along the top of a second, thicker gel (1.5 mm 12.5%) and sealed with 1% (w/v) agarose. The gel was then run at 20 mA and fixed in 10% (v/v) acetic acid/10% (v/v) methanol. The gels were dried on a DryGelSr for 1 h 10 min at 80°C and the markers were spotted with ^{35}S met (diluted 1/500000). The gel was then exposed to a phosphorimage plate overnight and developed on a FL-7000 phosphorimager.

2.3.4 Semi-permeabilised (SP) cell assay

SP cell generation uses a detergent to permeabilise the cell membrane whilst keeping organelle membranes such as the ER intact (Wilson, Allen et al., 1995). Cells co-transfected with A33 IgG4 HC and GFP and un-transfected cells used as a control were each prepared in two separate 15 cm dishes. Each dish was starved in minus cysteine and minus methionine (-cys-met) DMEM medium for 30 min and then replaced with fresh -cys-met containing radiolabeled methionine and cysteine mix (^{35}S) for 30 min at a concentration of 110 $\mu\text{Ci/ml}$. The cells were washed with PBS containing 20 mM NEM and trypsinised with 5ml 0.05% Trypsin-EDTA. Once the cells had detached, 8 ml of KHM buffer (20 mM HEPES pH 7.2, 110 mM potassium acetate, 2 mM magnesium acetate containing 100 $\mu\text{g/ml}$ soybean trypsin inhibitor) was added to the cell suspension. All samples were kept at 4°C throughout the protocol. The cells were then pelleted by centrifugation at 250 xg for 3 min at 4°C. The cell pellet was then re-suspended in 1ml ice-

cold KHM buffer (20 mM HEPES pH 7.2, 110 mM potassium acetate, 2 mM magnesium acetate containing 40 µg/ml digitonin) and incubated on ice for 5 min. To test that the cells had been successfully permeabilised, 10 µL of the sample was incubated with 10 µL 0.4% (w/v) trypan blue in PBS and the live cells were counted using a haemocytometer, if >90% were blue then the protocol was continued. The SP cells were isolated using centrifugation (250 x g for 3 min at 4°C) and the supernatant was kept as the cytosolic fraction. The SP cells were washed in 1ml KHM buffer and centrifuged (250 x g for 3 min at 4°C) and then resuspended in 1 ml HEPES buffer (50 mM HEPES pH 7.2, 90 mM potassium acetate) and incubated on ice for 10 min. The sample was centrifuged (250 x g for 3 min at 4°C) again and the supernatant of the HEPES fraction was collected as each fraction was analysed for confirmation of successful separation of the cytosol and the ER. The SP cell pellet was then resuspended in 120 µL of cold lysis buffer (50 mM TrisHCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 1% (v/v) Triton X-100 and 20 mM NEM). All fractions were affinity purified using protein A Sepharose beads before gel electrophoresis. The supernatant of each sample post affinity purification was kept and immunoprecipitated using GFP Trap (chromotek) following the manufacturers instructions.

The affinity purified samples were loaded onto two gradient (4-20%) polyacrylamide gels; one was developed using a FL-7000 phosphorimager (as described in section 2.3.2) and the other was used for western blot analysis (as described in section 2.3.1).

The immunoprecipitated GFP trap samples were loaded onto a gradient (4-20%) polyacrylamide gel and developed using a FL-7000 phosphorimager (as described in section 2.3.2).

2.3.5 Trichloroacetic acid (TCA) precipitation of the cell medium

CHO cells were co-transfected with various heavy chains plus $\alpha 1$ anti trypsin and incubated for 48 h. The medium was isolated from the cells by centrifugation at 250 xg for 5 min. A total volume of 12 ml of medium was collected, from this 1.2 ml was used and incubated with 300 µL of 100% (w/v) TCA (Sigma-Aldrich) (final concentration of 20% TCA) and vortexed. The sample mix was then incubated at 4°C for 30 min to allow the proteins to precipitate. The sample was then centrifuged at 21400 xg for 15 min. The protein pellet was then washed with 300 µL ice-cold acetone and centrifuged at 21400xg for 30 min. This process was repeated two times. The supernatant was removed, the pellet was dried and dissolved in 20 µL 2 X sample buffer. Any acidic samples were neutralised with 2 µL 1 M Tris.

2.3.6 Silver stain

Polyacrylamide gels post gel electrophoresis was fixed in a solution of 50% (v/v) methanol, 12% (v/v) acetic acid and 0.05% (v/v) formalin for 10 min. The gel was then washed using a solution of

35% (v/v) ethanol three times for 5 min. The gel was then sensitized in a solution of 0.02% (w/v) $\text{Na}_2\text{S}_2\text{O}_3$ and washed with H_2O three times for 5 min. The gel was then stained using a solution of 0.2% (w/v) AgNO_3 and 0.076% (v/v) formalin for 20 min, and washed with H_2O two times for 1 min. The gel was developed in 6% (w/v) Na_2CO_3 , 0.05% (v/v) formalin and 0.0004% (w/v) $\text{Na}_2\text{S}_2\text{O}_3$ until protein bands were visualised on the gel. Once the gel was appropriately developed, the reaction was stopped using a solution of 50% (v/v) methanol and 12% (v/v) acetic acid for 5 min. The gel was kept at room temperature in H_2O for storage.

2.3.7 Preparation of protein samples for mass spectrometry using a trypsin digest

Once the polyacrylamide gel containing the isolated transfected HC A33 had been silver stained and a band of interest was identified, a gel slice was cut of approximately 1mm X 2mm, transferred to an eppendorf and further cut into 4-5 pieces (another piece identical in size and position of the gel was also cut from the un-transfected negative control). The gel slice was prepared for trypsin digest by incubating with a solution of 50% acetonitrile and 50% 25 mM ammonium bicarbonate for 30 min. The protein within the gel pieces was then reduced using 10 mM DTT in 25 mM ammonium bicarbonate for 1 h at 56°C. To quench any free thiols within the protein, the sample was then treated with 50 mM iodoacetic acid in the dark for 45 min. The gel was then dehydrated via vacuum centrifugation at 20°C. Trypsin was added to digest the protein overnight.

Once the sample was prepared, it was sent to William Mullen at the University of Glasgow to carry out mass spectrometry under the following conditions. Tryptic digests were analysed on a Dionex Ultimate 3000 RSLC nano flow system (Dionex, Camberly UK). After loading (5 μl) onto a Dionex 0.1x20 mm 5 μm C18 nano trap column at a flowrate of 5 $\mu\text{l}/\text{min}$ in 98% 0.1% formic acid and 2% acetonitrile, sample was eluted onto an Acclaim PepMap C18 nano column 75 μm x50 cm, 2 μm 100 Å at a flow rate of 0.3 $\mu\text{l}/\text{min}$. The trap and nano flow column were maintained at 35°C. The samples were eluted with a gradient of solvent A:98% 0.1% formic acid, 2% acetonitrile verses solvent B: 80% acetonitrile, 20% 0.1% formic acid starting at 1% B for 5 minutes rising to 25% B after 45 min and finally to 50%B after 60 min. The column was then washed and re-equilibrated prior to the next injection. The eluant was ionized using a Proxeon nano spray ESI source operating in positive ion mode into an Orbitrap Velos FTMS (Thermo Finnigan, Bremen, Germany). Ionization voltage was 2.6 kV and the capillary temperature was 250°C. The mass spectrometer was operated in MS/MS mode scanning from 380 to 2000 amu. The top 20 multiply charged ions were selected from each scan for MS/MS analysis using CID at 40% collision energy. The resolution of ions in MS1 was 60,000 and 7,500 for CID MS2. MS and MS/MS data files were

searched, in this case, against the IPI Human non-redundant database using SEQUEST (by using Thermo Proteome Discoverer), with trypsin as enzyme specificity. Peptide data were extracted using high peptide confidence and top one peptide rank filters. Set Carboxymethyl of cysteine (C) as static modification and Oxidation of Methionine and Proline (M, P) as variable modifications and a peptide mass tolerance of ± 10 ppm and a fragment mass tolerance of ± 0.8 Da and allow for a maximum of two missed cleavage. Determine the false discovery rates by reverse database searches and empirical analyses of the distributions of mass deviation, whereby Ion Scores can be used to establish score and mass accuracy filters (protocol provided by William Mullen, University of Glasgow).

2.3.8 Protein digests using Endo H and PNGase F

A33 HC IgG4 was transfected into 24 million viable CHO-S cells using NovaCHOice. After 48h incubation post transfection cells were split into three samples and the media was separated from cells by centrifugation at 302 x g for 5 min. Cell lysates and media were prepared as described in section 2.3.1 and then affinity purified using protein A Sepharose. Protein A Sepharose beads were washed three times with immunoisolation buffer (50mM TrisHCl pH8, 1% Triton X-100, 150 mM NaCl, 2mM EDTA, 0.5 mM PMSF) then re-suspended in 50 μ L of 1X denaturing buffer (0.5% SDS, 0.04 M DTT, New England BioLabs). Each sample was boiled for 10 min and centrifuged at 16162 x g for 1 min. 2.5 μ L G5 buffer (0.5M Sodium Citrate pH 5.5 New England BioLabs) and 1000 units Endo H were added to one lysate and media sample, 2 μ L G7 buffer (0.5 M Sodium Phosphate pH 7.5 New England BioLabs), 2 μ L 10% NP40 and 1000 units PNGase F were added to a separate lysate and media sample and the third samples were incubated in buffer alone as a control (Freeze & Kranz, 2010). All samples were incubated at 37°C overnight. 4X SB was added to each protein sample the next day which were separated by SDS-PAGE on a gradient (4-20%) polyacrylamide gel. The gel was then silver stained for development (see section 2.3.6). If only one enzyme was used the protocol remained the same except for equally dividing into two parts instead of three.

2.4 Protein expression and purification

2.4.1 Expression and Induction of Protein using Isopropyl β -D-1-thiogalactopyranoside (IPTG)

BL21 Star (DE3) pLysS competent cells were transformed with either A33 VH domain, CH1 domain or CL domain in a pET 28a vector. Positive colonies were selected on kanamycin antibiotic agar plates after 16-18 h. The colonies were grown in 10 ml of Luria broth in the presence of kanamycin (50 μ g/ml) until they reached an OD of 0.5 at 37°C. Each culture was then induced

with 1 mM IPTG for 3h. Prior to induction, 1 ml of each culture was centrifuged at 144 xg for 1 min to pellet the cells representing the whole cell fraction for SDS PAGE analysis. Post induction samples were collected at a time point of 3 h. To normalise the samples, appropriate amounts of 2 X sample buffer (100 mM TrisHCl pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol) were added to each sample to correspond with the OD (OD 0.5 = 50 µL). Each sample was loaded on a 15% SDS PAGE gel for resolution of the protein of interest and Coomassie blue (10% (v/v) phosphoric acid, 10% (v/v) ammonium sulphate, 0.12% (w/v) Coomassie G250, and 20% (v/v) methanol) stained for visualisation of protein bands. A glycerol stock was made of positive cultures which expressed the protein of interest which were used for future cultures for protein purification.

2.4.2 Purification of the CH1 Domain from Inclusion bodies

BL21 Star (DE3) pLysS *E.coli* expressing the CH1 domain was used to inoculate a 25 ml culture at 37°C overnight. The following day the cells were sub-cultured into 1 litre fresh LB and grown until the culture had reached an OD of 0.5; the cell suspension was then induced with 1 mM IPTG. The culture was grown for 3h post induction and then pelleted at 240 xg for 15 min. The pellet was re-suspended in 20 ml lysis buffer (50 mM Tris pH8, 150 mM NaCl, 1% (v/v) Triton X 100, 10 µg/ml DNase, 100 µg/ml lysozyme, 1 tablet EDTA free protease inhibitor) and incubated for 10 min at 4°C. The cell pellet was then frozen in liquid nitrogen and thawed at 37°C; this process was then repeated four times. The cell lysate was incubated at room temperature for 10 min and then centrifuged at 2300 xg for 30 min. The cell pellet was then re-suspended in 20 ml lysis buffer plus 10 mM DTT. The lysate was transferred to a hand-held homogeniser and pestle where the pellet was completely re-solubilised by grinding the homogeniser in a circular mechanism. The sample was then centrifuged at 1000 x g for 15 min. This process was repeated a further three times. For the final wash, phosphate buffered saline (PBS) was used instead to remove the Triton X 100 and to introduce phosphate to the sample as the re-solubilisation buffer contains phosphate. The washed inclusion bodies were then re-suspended in 20 ml re-solubilisation buffer (6 M Guanidine hydrochloride, 10 mM β-mercaptoethanol (BME) and 20 mM Imidazole pH 7.6 50 mM phosphate buffer) and incubated overnight. The next day the lysate was centrifuged at 2300 xg for 30 min and filtered using a 0.45 µm Minisart filter unit (satorius).

All purification experiments were carried out using the ÄKTA purifier (GE healthcare). Purification of the CH1 domain was performed on a 5ml HisTrap HP (GE healthcare) column. Firstly the lysate was loaded onto the 5ml HisTrap HP column to allow the protein to bind to the column. Then the guanidine hydrochloride was slowly removed and exchanged with 50 mM Tris pH8, 200 mM NaCl, 20 mM Imidazole at a flow rate of 1ml/min to allow the protein to re-fold. The protein was then

eluted with 50 mM Tris pH8, 200 mM NaCl and 500 mM Imidazole using a linear gradient up to 100% of the elution buffer over 6 column volumes (30 ml). Fractions were analyzed on 15% SDS PAGE gel and visualised using Coomassie blue staining. Fractions which contained protein of interest were pooled and concentrated using a vivaspin turbo 15 column (5,000 MWCO Generon). A buffer exchange was then carried out using the gravity protocol according to the manufacturer's instructions on a PD-10 desalting column (GE Healthcare) in 50 mM Tris, 50 mM NaF and 0.5 mM EDTA pH 7.45.

2.4.3 Purification of CL and A33 variable region domains

Cells taken from the glycerol stock of BL21 Star (DE3) pLysS *E.coli* which expressed the isolated CL and A33 VH domain proteins were used to inoculate separate 25 ml cultures at 37°C and were grown overnight. The next day they were sub-cultured into 1 litre of fresh LB and grown until each culture had reached an OD of 0.5. Cultures were then induced with 0.1 mM IPTG and grown overnight at 16°C. The following day cultures were then pelleted at 240 xg for 15 min. The pellets were re-suspended in 20 ml lysis buffer (50 mM Tris pH8, 150 mM NaCl, 1% (v/v) Triton X 100, 10 µg/ml DNase, 100 µg/ml lysozyme, 1 tablet EDTA free protease inhibitor) and incubated for 10 min at 4°C. The cell pellets were then frozen in liquid nitrogen and thawed at 37°C three times. The cell lysates were incubated at room temperature for 10 min and then centrifuged at 2400 xg for 30 min. Each sample was filtered using a 0.45 µm Minisart filter unit (satorius) and subsequently loaded onto a 5 ml HisTrap HP (GE healthcare) column equilibrated in 50 mM Tris pH8, 200 mM NaCl, 20 mM Imidazole and eluted in the same way as the CH1 domain. Fractions were analyzed in the same manner as previously described and the pooled samples were concentrated using vivaspin turbo 4 columns (3,000 MWCO Generon). The samples were concentrated to a volume of 1 ml which was applied to a Superdex 200 10 300 (GF) gel filtration column (GE healthcare) equilibrated in 50 mM Tris, 50 mM NaF 0.5 mM, EDTA pH7.45 at a flow rate of 0.5 ml/min. Samples were again analyzed on a 15% SDS Page gel with Coomassie blue stain. All protein concentrations were measured using Cary 300 Bio UV-Visible Spectrophotometer (Varian).

2.4.4 Circular Dichroism (CD)

A Jasco J-810 spectropolarimeter was used for all CD measurements which were carried out at 20°C. All spectra were corrected for protein concentration and cell pathlength. The experimental procedure and calculations were conducted by Dr Sharon Kelly. Far UV spectra were measured in a 0.02cm pathlength quartz cuvette and recorded from wavelengths of 190 to 250 nm. An average of 3 spectra was recorded for each sample at a scan rate of 10 nm/min using a 2 sec response time and a 1 nm bandwidth. Spectra of the isolated domains were carried out at

concentrations of 1.349 mg/ml A33, 0.71 mg/ml CL and 0.5 or 0.77 mg/ml CH1. Purified protein samples were first incubated at 25°C for 4 h either individually or as a mixture at a ratio of 3:1 with the A33 or CL domain being in excess of the CH1 prior to CD measurements which were taken (Feige et al., 2009). The spectrum of the CH1 domain when incubated with either the CL or A33 VH domain was determined by comparing the combined sample spectra to that of the theoretical spectra. The theoretical spectra were determined by the addition of the spectrum of each domain in isolation. The spectrum of the isolated CH1 domain from the spectrum of the mixed domains was determined by subtracting the isolated spectrum of the CL domain or A33 domain from the spectrum of the domains mixed together.

Chapter 3 : An investigation into the cleavage of the HC of IgG4 with the variable domain A33.

3.1 Introduction

There are five classes of immunoglobulin (M, G, A, D and E), the simplest of which, and the focus of this investigation is IgG. IgG is a symmetrical molecule consisting of two identical LCs and two identical HCs which are held together via interchain disulphide bonds between the CH1 and CL domains of either chain (Liu & May, 2012). The HC consists of three constant domains and one variable domain, whereas the LC consists of one constant and one variable domain. When all chains are bound together the overall quaternary structure gives the antibody a “Y” shape (Figure 3.1) (Deisenhofer et al., 1976a, Deisenhofer et al., 1976b). The constant domains are named as such as these are identical amongst the same isotypes of IgG; however the variable domains differ with each antibody as this is the part of the protein which recognises foreign antigens (Heads, Adams et al., 2012). There are four isotypes of IgG (IgG1-4) which differ in the constant domain sequence of the HC and their disulphide bond arrangement between the LCs and HCs at the hinge region. For the purposes of this investigation, IgG1 and IgG4 have been studied as these immunoglobulins are the class which are used for therapeutics in industry. The greatest point of variability between the two isotypes is the hinge region between the Fab and Fc fragments which not only differs in length (IgG1 is longer than IgG4) but also sequence and the position of the interchain disulphide between the CL domain of the LC and the CH1 domain of the HC (Aalberse & Schuurman, 2002). The hinge region of the antibody is also the most vulnerable and flexible part of the protein giving the advantage of movement when conducting the functionality of IgG via binding to its target antigen, this also differs between IgG1 and IgG4; IgG1 being more flexible when comparing the hinge-folding mode of flexibility (Roux, Strelets et al., 1997).

There have been several assays that have been used to investigate the stability of immunoglobulins (Correia, 2010). The most well-established of these assays is the protein's susceptibility to enzyme cleavage. As the hinge region is the most vulnerable, it is this part of the protein which is most sensitive to cleavage by proteolytic enzymes such as pepsin. Cleavage with pepsin separates the protein into a Fab dimer and a single Fc fragment as it cleaves below the hinge region (Turner et al., 1970). Another example is papain, which cleaves immunoglobulin at the site SCNKTHT splitting IgG into single Fab domains and the Fc fragment as it cleaves above the hinge region and is how these fragments were first identified (Jefferis et al., 1968) (Cordoba, Shyong et al., 2005). Papain is a cysteine specific protease which is why it attacks above the hinge region using the first of the interchain disulphide bonds in its nucleophilic attack. There are four main groups of protease enzymes all of which generate a nucleophile that attacks a peptide carbonyl group at specific sites within a polypeptide chain (Berg et al., 2006). These protease groups involve either the amino acids serine, cysteine, aspartic acid or metal ions (metalloproteases) to insight their nucleophilic attack. The most common metal involved in the

action of metalloproteases is Zinc where the enzyme uses the metal to protonate a water molecule in the action of attack. In some cases more than one metal ion can be involved, for example metallo- β -lactamase requires the action of two Zinc ions for its function to break down a β -lactam ring present in most antibiotics (Mitić, Miraula et al., 2014). These proteases are one of the major contributors to antibiotic resistance. Secreted metalloproteases from bacteria have been shown to cleave avian IgG as a form of attack to inhibit the host's immunity (García-Gómez, Vaca et al., 2005). Therefore understanding the susceptibility of immunoglobulin to cleavage via a wide range of proteases is important to protect against the adapted attack by pathogens.

Understanding of antibody folding and assembly initially involved identifying the components and domains of the antibody structure to investigate how they are arranged together. Previous investigations on antibody folding and assembly have used the approach of pulse chase experiments while over-expressing ER resident chaperones which are involved in antibody assembly (Ailor & Betenbaugh, 1998, Hsu, Watson et al., 1996). The folding and assembly of several antibodies have been investigated and are described throughout this chapter. One observation of these investigations was the cleavage of one particular HC, which was specific to this construct. This chapter outlines and investigates the cause and point of cleavage of this particular HC.

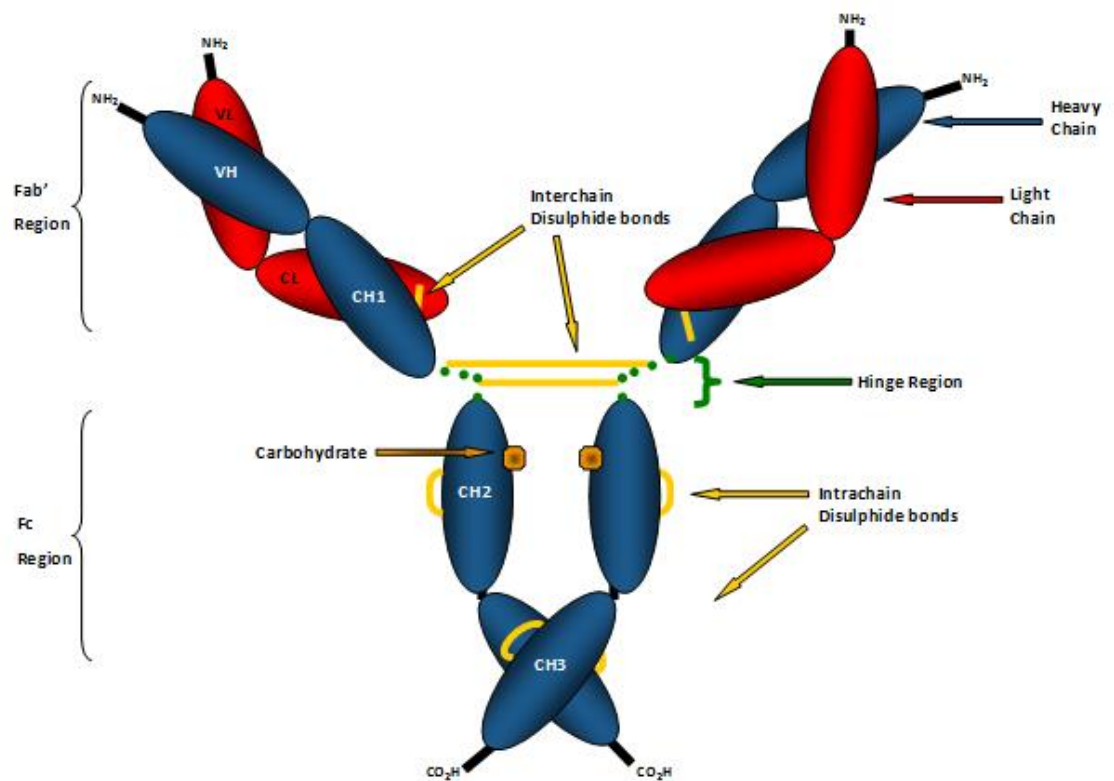


Figure 3.1: Schematic structure of IgG and each domain.

IgG is comprised of two identical LCs (red) and two identical HCs (blue) which are held together via interchain disulphide bonds. The Fab fragment of the antibody consists of the LC and the VH and CH1 domains of the HC. The Fc fragment consists of the CH2 and CH3 domains of the HC connected via the disulphide bridge at the hinge region. The hinge region of the antibody connects the Fab and Fc fragments and is the most vulnerable area of the protein.

3.2 Aim

Using the initial approach of pulse chase experiments, the folding and assembly of the antibody A33 IgG was investigated. This antibody is denoted as A33 as this is the name of the antigen it recognises (Heath, White et al., 1997). These investigations will also try and interpret the cause of the cleavage of this antibody of the isotype IgG4 and why this finding is unique to this construct.

3.3 Identification of a cleavage product of A33 IgG4

Pulse chase analysis of various immunoglobulins was carried out to identify the intermediates formed during the folding of IgG. Initially, two isotypes of IgG (IgG1 and IgG4) with the same variable domain of A33 were investigated to note any difference between their assembly. CHO cells were co-transfected with A33 κ LC and either A33 IgG4P HC or A33 IgG1 HC. For the purposes of this chapter A33 IgG4P is denoted with a "P" after its name to note a point mutation within the hinge region. The wild type of IgG4 contains the sequence CPSC which causes half molecule formation after secretion, however, this does not occur with IgG1 as it contains the sequence CPPC within its hinge region, the proline instead of a serine aids in stabilising the two interchain disulphide bonds at each cysteine. Half molecule formation is when the two complete antibody chains containing both the HC and LC dissociate in solution after secretion and form two halves of IgG each consisting of a full HC with the Fab and Fc fragment and a single LC. All IgG4 antibodies described in this thesis have the S225P mutation within its hinge region to avoid half molecule formation unless stated otherwise.

CHO cells were transfected with A33 IgG1 and A33 IgG4P for a 48 h incubation period. The cells were then radiolabelled with ^{35}S cysteine/methionine containing medium for 30 minutes and chased for a total of 1 hour with time points at 0, 5, 30 and 60 minutes. The cells were treated with 20mM NEM to freeze the disulphide status of the proteins and then lysed and immunisolated with anti κ LC and anti Fc HC in the presence of protein A Sepharose beads to isolate any folding intermediates generated during IgG assembly. Immunoglobulin is known to bind to protein A via the Fc region of the antibody and is, therefore, a useful tool in isolating antibodies from cellular lysates and medium (Sjodahl, 1977, Sjodahl, 1977). The isolated radiolabelled proteins were visualised on a gradient (4-20%) SDS PAGE polyacrylamide gel under non-reducing conditions (Figure 3.2A+B). A reduced sample in the first lane of each gel was also included to observe which multimer species of immunoglobulin were present and a cell lysate transfected with an empty vector negative control was included to observe the protein bands which were specific to IgG. The banding patterns produced from IgG1 and IgG4 are largely similar; both containing complete IgG with two LCs and two HCs (approximately 200kDa), an intermediate migrating slightly faster representing a HC dimer with a single LC (approximately 150

kDa) and protein band representing a HC dimer migrating faster still (approximately 135 kDa) which contains no LC (identification of proteins which make up each band explained further in Chapter 4). All of these complexes are dissociated and are no longer observed upon addition of DTT (reduced lane Figure 3.2A+B) confirming that they are held together via disulphide bonds. One striking difference between the two banding patterns, however, was the presence of a protein band which migrated to approximately 30kDa which was present for IgG4P A33 but not IgG1. The band in question is also unchanged over the course of the hour indicating that it exists independently and is not being secreted or degraded (protein band is not present in the media – Chapter 4 Figure 4.2A).

In order to determine the identity of the protein in question, CHO cells were again transfected with A33 IgG4P and immunisolated as described earlier. The protein sample underwent gel electrophoresis under non-reducing conditions with a transfected empty vector as a negative control, the polyacrylamide gel was then silver stained to visualise the protein bands (Figure 3.3B). The 30kDa protein band of interest plus an equivalent size of gel migrating to the same place in the empty vector negative control were excised from the gel and underwent a trypsin digest to prepare the sample for mass spectrometry analysis. The peptides identified from this analysis corresponded to the CH2 and CH3 domains of IgG4 indicating that this portion of the antibody is being cleaved (Figure 3.3A and C). As the size of the protein band in question is approximately 30kDa, this implies that the Fc fragment is lacking the interchain disulphide formed through the hinge region. However, whether the cleaved product exists as a dimer or monomer in solution is unclear as it may form a dimer via non-covalent interactions which would be abolished when the proteins are denatured and run on a polyacrylamide gel. This conclusion is supported by the fact that the protein band remains unchanged whether it is reduced or non-reduced indicating there are no disulphides being formed between the two Fc fragments of the HC (Figure 3.2A).

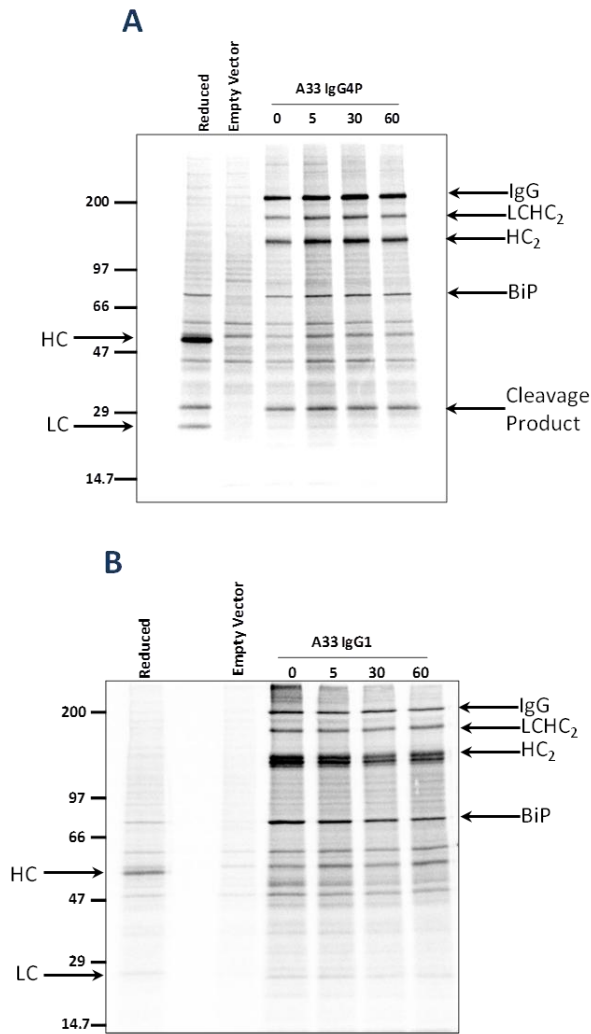


Figure 3.2: Pulse chase analysis of IgG1 and IgG4 with the variable region A33.

60 minute pulse chase of A33 IgG4P (A) and A33 IgG1 (B). CHO cells were transfected with equal amounts of LC and HC. Both experiments include a reduced sample and a transfected empty vector negative control after 60 minutes of chase time. The transfected samples were analysed at time points of 0, 5, 30 and 60 minutes and purified using anti κ LC, anti Fc HC and protein A Sepharose beads. Several intermediates can be identified from the time course including: IgG, LCHC₂, HC₂ and BiP. In the reducing lane these intermediates have lost their disulphide bonds and therefore are identified as a single LC and HC. There is an additional band present for the IgG4P pulse chase which migrates to approximately 30 kDa which has been identified as a cleavage product of the transfected HC.

3.4 Classification of the cleavage product of A33 IgG4

The three peptides corresponding to amino acid sequences within the CH2 and CH3 domains of IgG4 identified by mass spectrometry align to sequences C-terminal to the hinge region including the glycosylation site within the CH2 domain. The data from the mass spectrometry show that there is an asparagine residue within the glycosylation site; this information suggests that the protein has not been glycosylated. This is because if glycosylation had occurred, the asparagine in the peptide analysis would have either appeared as glycosylated, or if it had been de-glycosylated it would have been converted to aspartate instead. Therefore, to test the glycosylation status of the cleavage product of IgG4 the lysate sample of the transfected cell was incubated with Endo H. Endo H is an enzyme which cleaves asparagine linked mannose rich oligosaccharides. When glycoproteins enter the ER to fold they are glycosylated via an asparagine amino acid within a glycosylation site, therefore Endo H will cleave glycosylated proteins which exist in the ER.

CHO cells were transfected with A33 IgG4P and affinity isolated using protein A Sepharose. The protein samples were denatured, reduced and incubated with either Endo H or buffer as a negative control. The samples were then run on a gradient (4-20%) polyacrylamide gel and silver stained to visualise the protein bands (Figure 3.3D). Upon Endo H treatment it can be seen that the full length HC which has been effectively de-glycosylated has a faster mobility through the gel after losing its carbohydrate indicating the enzyme has successfully cleaved the protein. However, it can be seen that the mobility of the cleavage product does not alter upon Endo H treatment meaning that no cleavage has occurred; therefore, suggesting the protein is not glycosylated. This was unexpected as there is not a clear reason why the cleavage product should not be glycosylated as it contains a glycosylation site and the full length un-cleaved HC has shown to be glycosylated.

One possibility as to why we are not seeing glycosylation of the cleavage product is that it could actually be a mislocalised product meaning it would be translated and exist in the cytosol and never actually enter the ER for glycosylation to occur. This can be caused by an internal initiation event within the polypeptide chain. Alternative initiation and splicing is a common event for some proteins, for example p53 which is a transcription factor with 12 different variations (Marcel, Perrier et al., 2010). In the case of the cleavage product, an alternative AUG start codon could be used as an initiation of translation causing the mislocalisation of the Fc portion of the immunoglobulin polypeptide chain. In order to test this hypothesis the cytosol and the ER compartments of the cell were separated. CHO cells were co-transfected with the HC of IgG4P A33 and GFP which is translated and expressed in the cytosol of the cell. Therefore, when the cell is fractionated the GFP should only be present in the cytosolic fractions. The transfected cells

were radiolabelled with ^{35}S cysteine/methionine media for 30 minutes. Prior to cell lysis the cytosol was released by preparing semi-permeabilised (SP) cells. This preparation involves treating the cells with a cholesterol-specific detergent digitonin which permeabilises the cell membrane, releasing the cytosol while the ER remains intact. Several fractions were collected throughout this process; the digitonin fraction (which represents the cytosol), a HEPES buffer wash fraction (which also represents proteins in the cytosol) and the final SP cell (which represents the ER) (successful separation of the cytosol and ER can be seen in Appendix 3). Each fraction was collected for both the transfected cells and an un-transfected negative control. CHO cells transfected with HC A33 IgG4P which remained intact were also used in this experiment as a positive control (Figure 3.3E). All samples were affinity purified using protein A Sepharose and run on a gradient (4-20%) polyacrylamide gel. After the affinity purification of the antibody products the samples were purified again using GFP-trap protocol (described in Chapter 2, section 2.3.4) to isolate the GFP co-expressed with the HC (Figure 3.3E panel 2). Upon analysis of the radiolabelled proteins it can be seen that the cleavage product of HC IgG4P A33 is present in the SP cell prep and is not present in either the digitonin or HEPES wash steps which in this case represent the cytosol. It can also be seen that the GFP is present in the cytosol fractions further confirming that the cytosol and ER have successfully been separated. The un-transfected negative control does not contain any protein migrating to 30kDa whereas the transfection of HC A33 IgG4P intact cell positive control does. This is strong evidence suggesting that the cleavage product of HC IgG4 A33 is present in the ER with the full length HC and is unlikely to be a mislocalised product. Furthermore, observing the peptide data from the mass spectrometry analysis the only methionine which could be perceived as a start codon to initiate mislocalisation is within the first peptide (Figure 3.3A: represented by a yellow arrow). As there are peptides which were recognised from the mass spec data upstream of this methionine, it suggests that an event of alternative initiation may not have occurred. This strengthens the argument that this cleavage product is generated within the ER and not a mistranslation product in the cytosol. Therefore, why and how cleavage is occurring, and why glycosylation is not are important questions which need to be addressed.

3.5 Investigation of IgG4 mutants to address why the cleavage product is not seen with IgG1

As described earlier, the cleavage product is seen with A33 IgG4P but not with A33 IgG1. Both HCs have an identical variable domain, however, the sequence of the constant regions differ between each isotype. To address where cleavage is occurring, the most logical area of the HC to investigate is the hinge region as this is the most vulnerable area to digestion of the full length HC. Furthermore there were no peptides recognising sequences corresponding to the hinge region from the mass spectrometry data indicating that this is potentially the area where cleavage is occurring. Furthermore, the HC of A33 IgG4 has a point mutation within the hinge region S225P which could be affecting cleavage of the protein. To investigate if this was the case, the WT HC of IgG4 A33 was generated with the sequence CPSC instead of CPPC at the disulphide bridge within the hinge region. CHO cells were transfected with A33 κ LC and A33 IgG4 WT HC and incubated for 48 hours. The transfected cells were then radiolabelled with ^{35}S cysteine/methionine media for 30 minutes and prepared as previously described (Figure 3.4C, IgG1 and IgG4P results were also included in this figure as a positive and negative comparison (Figure 3.4A +B)). Separate cells transfected with an empty vector were included as a negative control and to identify which protein intermediates were specific to IgG. The cleavage product was still present for A33 IgG4 WT. Therefore, the S225P mutation was not causing the effect of cleavage.

There are 5 differences in the sequences between the hinge region of IgG1 and IgG4, and an extra four amino acids within the IgG1 protein which lengthen it. These differences also contribute to the difference in the disulphide bond arrangement between the HC and LC. As the cleavage product is seen with IgG4 and not IgG1, a mutant which contained these differences was created. Therefore the VH, CH2 and CH3 domains still remain as IgG4, but the hinge region of the HC now resembles that of IgG1. If cleavage is occurring around the hinge region of IgG4 and not IgG1 then the hypothesis would be that a HC IgG4 with the hinge region of IgG1 would no longer produce the cleavage product. This was tested using a radiolabelling mix of ^{35}S cysteine/ methionine as described previously (Figure 3.4D). Cells transfected with an empty vector negative control were also used, and both samples underwent gel electrophoresis under non-reducing conditions. Upon analysis however it can be seen that the cleavage product is still present for the A33 IgG4 IgG1 hinge mutant. This result indicates that the cleavage product is not caused by the sequence differences within the hinge region; however the rest of the CH1 domain differences (of which there are 8 amino acid differences) have yet to be investigated to determine whether any of these subtle changes alter the fate of the expressed HC.

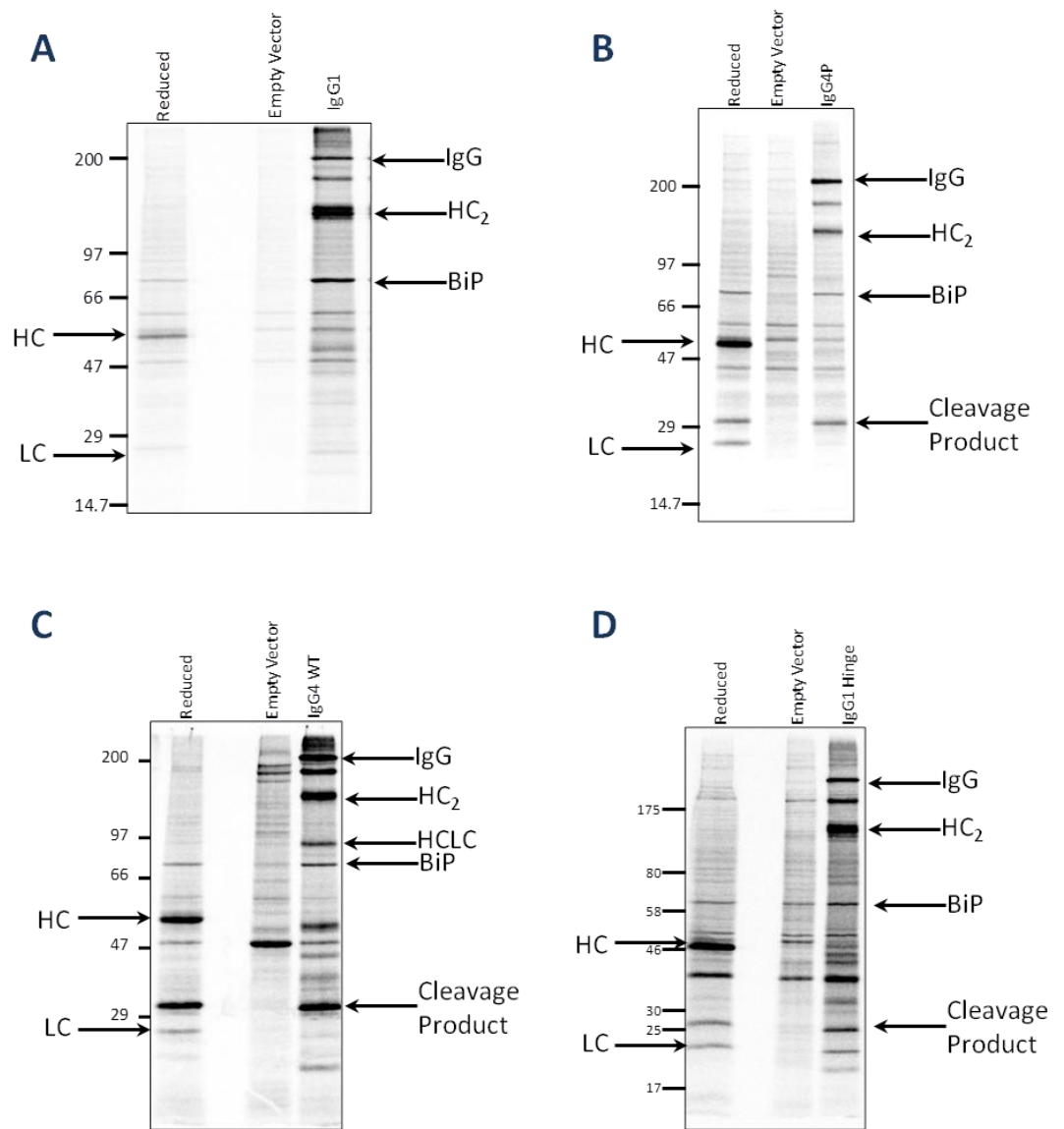


Figure 3.4: Comparisons of IgG1 and IgG4 constructs to identify evidence of the cleavage product.

CHO cells were co-transfected with equal amounts of LC and HC and radiolabelled for 30 min. Each gel consists of a reduced sample after 60 min of chase time, a transfected empty vector negative control and a non reduced sample after 30 min of pulse time. All samples are affinity purified crude lysate using protein A Sepharose. A: Co-transfection of LC and HC IgG1 (A33) B: Co-transfection of LC and HC IgG4P (A33) C: Co-transfection of LC and HC IgG4 (A33) D: Co-transfection of LC and HC IgG4P-IgG1 hinge (A33).

3.6 Cleavage of the HC A33 IgG4 only occurs with this construct

The reason for the presence of cleavage of the HC seen for A33 IgG4 and not IgG1 still remains unclear. However, another question to address regards whether or not cleavage of the HC is present with other IgG4 HCs. Six other HCs were investigated which all have the identical constant regions but only differ in their VH sequence. The first HC investigated 497 IgG4 HC is an antibody developed by UCB Pharma. The other five HCs are commercially available and for the purposes of this investigation have been denoted HC1-5. HCs1-5 VH domains were sub-cloned into the same vector (pMhg4PFL) used for A33 IgG4 expression with the restriction sites HindIII and XhoI. The 497 IgG4 HC was also in the same expression vector as it was provided by UCB Pharma. All seven HCs were transfected into CHO cells and incubated for 48 hours with an untransfected negative control also included. The cells were then radiolabelled with ³⁵S cysteine/methionine media for 30 minutes and affinity purified using protein A Sepharose. The samples were run under reducing and non-reducing conditions on a gradient (4-20%) polyacrylamide gel (Figure 3.5). Of the seven HCs, the only one which displayed the presence of the cleavage product still remains to be A33 IgG4. Therefore, cleavage of the HC appears to be unique to A33 IgG4 and, therefore, not a phenomenon displayed by other antibody chains.

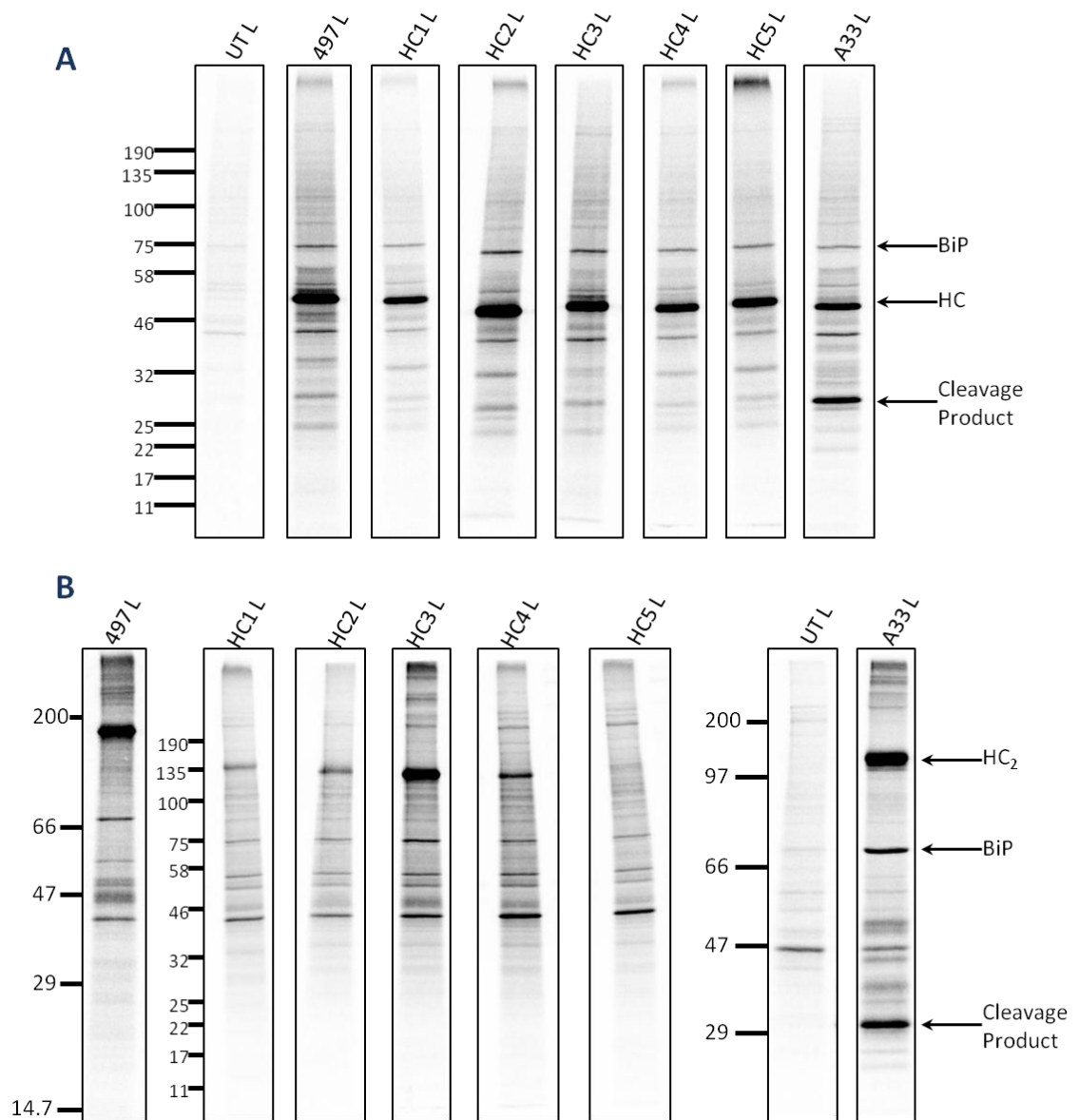


Figure 3.5: Screen of seven IgG4 HCs for presence of the cleavage product.

HCs were separately transfected into CHO cells using 20 μ g DNA. Each sample was radio-labelled for 30 min and affinity purified using protein A Sepharose beads. Each sample was run under reducing (A) and non-reducing (B) conditions. An un-transfected cell lysate was used as a negative control and A33 IgG4 which is known to produce the cleavage product was used as a positive control (last lane). Of the seven HCs, the protein with the VH domain of A33 undergoes cleavage to produce the cleavage product. Under reducing conditions, there appears to be minor evidence of the cleavage product but as this is not exhibited under non-reducing conditions it can be attributed to non-specific radiolabelled protein bands.

3.7 Discussion

Proteolytic cleavage of immunoglobulin fragments is a well-studied area of antibody structure and stability. The fragmentation which occurs is usually into the Fab and Fc counterparts of IgG which are separated via the hinge region. This part of IgG is less structured than the other domains and, therefore, more susceptible to cleavage. Several proteases have been reported to cleave this region either before (papain) or after (pepsin and matrix metalloproteinases (MMPs)) the hinge region (Brezski & Jordan, 2010). Interestingly, the MMPs investigated were shown to only cleave IgG (not other immunoglobulin HCs) in a two-stage process which included a bi-product of a 32kDa Fc fragment (Gearing, Thorpe et al., 2002). The suggested cleavage site by the MMPs also occurs upstream of the peptide data described in this chapter. However, the main target of these proteases are extracellular matrix proteins and are, therefore, unlikely to be active in the endoplasmic reticulum where the cleavage product is located, however, it does suggest that a similar protease may be causing cleavage of the A33 IgG4 HC. As the MMPs activity is reduced with the chelating agent EDTA, this suggests that if this enzyme is cleaving the HC then it is prior to cell lysis as that buffer contains EDTA plus a range of protease inhibitors. It has been noted that CHO cells do have the ability to secrete MMPs so this protease could be contributing to cleavage of the HC, however, if this protease was responsible, the specificity to A33 HC IgG4 would be unusual (Sandberg, Lutkemeyer et al., 2006). However, previous investigations on the susceptibility of different IgG isotypes to proteolytic enzymes has shown that IgG4 is cleaved more readily than IgG1 (Molla, Kagimoto et al., 1988). This difference in susceptibility could be the reason that the isotype IgG1 A33 seems to be immune to cleavage whereas IgG4 A33 does not. Another example of cleavage or degradation of the HC *in vivo* has been shown when the ER resident chaperone BiP was overexpressed in expression systems containing either LC and HC co-expressed together or HC alone. In both cases, the HC yielded a 35 kDa fragment after a 4 day incubation period which seemed to increase over time (Hsu, Eiden et al., 1994). This result highlights the predisposition and vulnerability immunoglobulins seem to have to degradation or cleavage by proteases.

Another protease which has been developed for its characteristics to cleave all isotypes of IgG is immunoglobulin-degrading enzyme of *Streptococcus pyogenes* (IdeS). In one study using IdeS the authors were able to cleave several antibodies varying in isotype into their Fab and Fc counterparts quickly to investigate different modifications which alter the properties of each IgG. One result which correlates with the data presented in this chapter was the use of the S228C mutation within the hinge region of IgG4 to avoid half molecule formation. This mutation yielded no difference in the ability of IdeS to cleave the antibody (An, Zhang et al., 2014). Also, it was noted that antibodies with different variable domains yielded different modifications. Another

investigation, which involved the V_HH domains of llamas digested in gastric juices to determine their suitability for use as an oral therapeutic, showed a varied level of stability to proteases when comparing seven different variable domains (Harmsen, van Solt et al., 2006). This implies that the stability of the variable domain could be contributing to the protein's overall susceptibility to cleavage.

Another potential advantage to understanding and identifying the different cleavage patterns of IgG is the role proteolysis plays in disease. Several parasites invade the body and attack the immune system using proteases which target immunoglobulins. For example, cruzipain a cysteine protease from the parasite *Trypanosoma cruzi* can cleave all IgG isotypes, again IgG1 and IgG4 varying in where they are cleaved and how they fractionate (IgG1: Fab and Fc, IgG4 Fab₂ and single Fc) (Berasain, Carmona et al., 2003). Another example which also cleaves all IgG isotypes each differing in the cleaved products produced is *Fasciola hepatica* (Berasain, Carmona et al., 2000). One common feature between both of these parasites and other protease studies described is that the cleavage of immunoglobulins seems to occur initially at the hinge region of the antibody. However, the evidence from this chapter has shown that altering the hinge region of IgG1 and IgG4 does not act in the prevention of cleavage of the IgG4 isotype. In order to aid in the prevention of these disease pathways an understanding of cleavage, but also protection of immunoglobulins is a vital area of research which needs to be explored.

One result demonstrated in this chapter is that the cleaved HC is not glycosylated. The lack of glycosylation to the HC may be causing or increasing the proteins susceptibility to cleavage. For example, in one study it was found that papain more readily cleaves IgG which has been deglycosylated. They found that prior removal of the glycans in the Fc region of the antibody resulted in a faster rate of papain-mediated degradation with a time of less than one hour compared to more than two hours when the HC was glycosylated (Raju & Scallan, 2006). This result also showed that a lack of glycans resulted in further degradation of the HC. This data correlates with the result that the cleavage product is not glycosylated and therefore is potentially more susceptible to cleavage. If the HC is being cleaved in the ER and cleavage is inhibited upon glycosylation, this could explain why there are two glycoforms of the HC present in the ER.

Furthermore, with the industry of the production and design of therapeutic antibodies growing rapidly the susceptibility and cause of cleavage and degradation of different immunoglobulins needs to be identified in order to improve production of the therapeutics. If a large percentage of the HC is being cleaved during the folding and assembly pathway, then the potential yield of the specific therapeutic would be decreased. In order to avoid this, it would be beneficial to the manufacturer if cleavage of the HC was initially tested, and if such an event occurred if the

cleavage could be inhibited the potential yield of the end product would be greatly increased as one would avoid the loss of the therapeutic protein to cleavage or degradation.

This chapter has identified a cleavage product specific to the antibody IgG4 A33. The 30 kDa fragment from the Fc portion of the HC is not glycosylated yet exists in the ER of the cell and is not secreted. This particular product was not seen with six other IgG4 HCs which had identical constant domains but different variable domains. It also was not seen with the isotype IgG1 which had an identical variable domain. It was still seen, however, with an IgG4 A33 HC which had the hinge sequence of IgG1. The overall conclusion is that the cleavage seems to be specific to the more unstable isotype (IgG4) with a contribution from the variable domain A33 giving this HC a higher susceptibility to proteolytic cleavage.

Chapter 4 : An investigation into the secretion of HC dimers in a LC independent manner

4.1 Introduction

How antibodies fold within the ER is a well-researched and characterised process. These investigations have led to the identification of several chaperones and oxidoreductases which aid and regulate the folding of proteins. Arguably, one of the most important of these discoveries was the chaperone BiP which plays an important role in the quality control of the ER and in the processes of antibody folding as it recognises hydrophobic residues which are exposed by unfolded proteins (Bole et al., 1986). The process of antibody folding begins with the HC and LC being co-translationally translocated into the ER where both chains begin to fold via internalisation of their hydrophobic residues. The LC has been shown to fold almost instantly upon entry into the ER as the first intrachain disulphide loop within the LC is formed in less than one second (Bergman & Kuehl, 1979). The HC forms a dimer before the addition of LC and being larger, also requires the aid of various chaperones to fold. The last domain to fold within the HC which is pivotal to the entire process of antibody folding is the CH1 domain. BiP binds to the exposed hydrophobic residues within the domain and, therefore, sequesters the HC dimer in the ER. The LC is also bound to BiP via its VL domain as the CL already folded does not associate; this is to prevent the LC being secreted without its partner HC (Davis et al., 1999, Hellman et al., 1999, Lee et al., 1999). The ability for BiP to bind to the VL domain, however, varies with each LC due to sequence differences between the VLs; it has also been shown that the ability of BiP to bind to a particular VL domain correlates with LC secretion and to the half-life of the LC (Knittler & Haas, 1992, Skowronek, Hendershot et al., 1998). However, this process of BiP quality control is not infallible as the LC has been shown to form dimers and be secreted separately and not as part of IgG (Mains & Sibley, 1983).

The folding of IgG continues as the CL domain of the LC displaces the chaperone BiP from the CH1 domain of the HC. The association of these two domains initiates folding of the CH1 domain as it has been shown that structure is introduced to the CH1 by the CL *in vitro* (Feige et al., 2009). Complete folding of the CH1 domain within the ER is also internally controlled by isomerisation of a conserved proline residue; the transition of the proline from a *trans* to a *cis* conformation which requires catalysis by a PPlase, enables the correct orientation of the chains for association with each other to allow disulphide formation (Lilie et al., 1995, Thies et al., 1999) (see Chapter 1 Figure 1.6).

The process of antibody assembly and secretion described above holds true for the majority of immunoglobulins. However, as mentioned earlier, the LC is able to sometimes escape the tightly regulated quality control of antibody folding, and this is also true for some mutants of the HC. For example nurse sharks have immunoglobulin/T cell receptor like proteins which consist of a single

V domain and five C domains which are secreted into serum absent from any LC-like protein (Greenberg et al., 1995). In addition, camels have also been shown to secrete HC dimers known as camelids, which are devoid of LC; however these HCs lack the CH1 domain, which as described is the domain which aids in the quality control of antibody folding and secretion in humans (Hamers-Casterman et al., 1993, Muyldermans et al., 1994). It has also been reported that some mutant IgMs were able to fold and be expressed on the surface of pre B cells lacking a surrogate LC (Minegishi & Conley, 2001). This was also demonstrated using a different mutant IgM which could no longer associate with surrogate LCs but was still expressed on the cell surface. These results question the mechanism of early B cell development, by highlighting that the surrogate LCs which no longer associate with the HCs have not inhibited their expression indicating they may not be necessary for the signalling function of pre B cells (Shaffer & Schlissel, 1997). Another study looking at altering ratios of LC and HC in stable CHO cell lines have noted the secretion of HC dimers but attributed this to aggregation and assumed secretion was due to overexpression implying the dimers were incorrectly folded (Ho et al., 2015). However, the investigations of antibody assembly and secretion discussed in this chapter have shown that some HCs have the ability to correctly fold and secrete HC dimers in a LC independent manner.

4.2 Aim

The overall aim of this chapter was to characterise and determine the mechanism underlining the secretion of HC dimers. The folding and assembly of single HCs were investigated along with isolated immunoglobulin domains to identify how the CH1 domain is able to fold within the HC in the absence of the CL domain.

4.3 HC dimer secretion occurs in the presence and absence of LC

Two isotypes of IgG (IgG1 and IgG4 A33) were investigated to observe any differences between the intermediates secreted by each antibody. CHO-S cells were transfected with equal amounts of A33 κ LC and either A33 IgG1 HC or A33 IgG4 HC (Figure 4.1 A). After transfection, the cells were incubated for 48 h to enable sufficient expression of the protein following which both the lysate and the media were collected. The antibodies were affinity purified using protein A Sepharose beads, denatured to release the protein from the beads and then run under non-reducing conditions on a gradient (4-20%) SDS PAGE gel. Detection of the antibody products was performed by western blotting using a fluorescently-labelled protein A which emits at 800 nm. The protein banding pattern of each isotope is similar for both the lysate and the media. There are three protein bands present in the lysate corresponding to LC₂HC₂, LCHC₂ and HC₂ as indicated (identification of each band described in section 4.4). As expected, secretion of IgG can be seen for both isotopes; however, there is also another protein band present in the media of IgG1 and

IgG4. The size of the protein band (approximately 100 kDa) is suggestive of a HC dimer. It is also of note that the protein band of HC dimer in the media has a faster migration through the gel when compared to its counterpart in the lysate. The non-reducing condition of the SDS PAGE analysis suggests that the faster migration of the HC is due to the formation of an intrachain disulphide bond possibly in the CH1. One way to test whether this band is a HC dimer secreted independent from LC was to transfect CHO-S cells with only HC to observe if secretion of the dimer occurred (Figure 4.1B). The HCs of IgG1 and IgG4 were separately transfected, isolated and developed in the same way as for the co-transfection with the LC. It can be seen that for both IgG1 and IgG4 HC dimers are secreted with the same migratory pattern of the HC dimers as seen for the co-transfection with LC, again implying the CH1 domain has formed an intrachain disulphide bond. This result demonstrates that HC dimer secretion is observed in the absence and presence of LC, indicating there is an alternative process to fold the CH1 domain other than the CL domain.

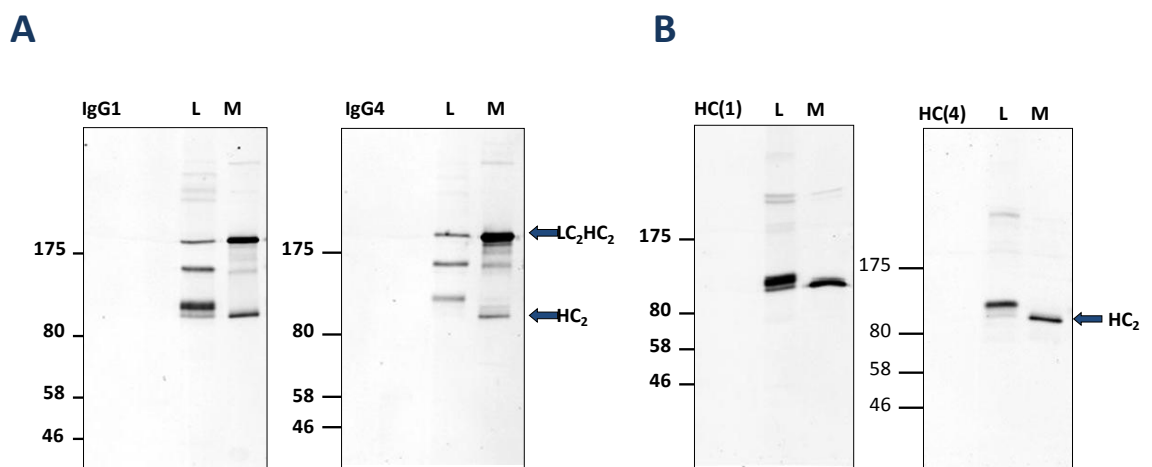


Figure 4.1: HC dimer secretion occurs in the absence and presence of LC.

CHO cells were co-transfected with equal amounts of LC and HC. Cell lysates and media samples were collected and prepared via Immunoisolation with protein A Sepharose. A & B: Western blots of IgG1 and IgG4 in the absence and presence of LC. L corresponds to intracellular material and M to extracellular secreted material.

4.4 Confirmation of HC dimer secretion

To confirm the initial result that HC dimers are actively secreted from the cell three approaches were taken; firstly the composition of the proteins of the antibody products were investigated, secondly the folding status of the CH1 domain was examined and thirdly their glycosylation status was investigated.

First, CHO cells were co-transfected with κ LC and IgG4 HC A33 or HC A33 alone and incubated for 48 h. The cells were then radiolabelled with ^{35}S cysteine/methionine media for 30 minutes and chased for up to 5 h with time points of 0, 1, 2, 3, 4 and 5 h (Figure 4.2 A + B). This time course was carried out to assess the optimum chase time to observe the contents of the antibody products secreted. A transfected empty vector negative control and reduced sample after 5 h were also included. It can be seen that over time, the IgG and HC dimer protein bands intensity is decreasing in the lysate and increasing in the media as seen at the 5 h time point (compare 5 h time points for lysate and medium Figure 4.2 A) where there is very little IgG present in the lysate anymore as the majority of radiolabelled protein now exists in the media. Another observation from the time course is that the ER resident chaperone BiP is also present and is, therefore, isolated via its association with the HC. As the secretion of HC dimers is occurring, the HC dissociates from BiP and is secreted from the cell in the absence of LC (Figure 4.2 C + D). From the result of this time course a time point of 3 h was deduced as the optimum to observe an equivalent radiation signal in both the lysate and the media samples.

To determine the protein content of each band, 2D gel electrophoresis was carried out. CHO cells were transfected and radiolabelled as described previously and chased for 3 h. The lysate and media were collected and affinity purified using protein A Sepharose beads. The lysate and media samples were denatured and run under non-reducing conditions as the first dimension on a gradient (4-20%) polyacrylamide gel with a width of 0.75 mm. Once the proteins had run through the first dimension, each lane was excised and incubated with DTT to reduce the proteins within the gel. They were then placed horizontally on top of a 15% polyacrylamide gel which was 1.5 mm thick and run in the second reducing dimension (Figure 4.2 E). There are three higher molecular weight bands within the lysate when run under non-reducing conditions, the top two are comprised of both LC and HC as seen by the reducing dimension (indicated by a red asterisk Figure 4.2 E lysate). From the molecular weight in the non-reducing dimension, it can be suggested that the first protein band is IgG (LC_2HC_2) and the second lower protein band corresponds to a folding intermediate comprised of a HC dimer with a single LC (LCHC_2). The third protein band (indicated by a blue asterisk Figure 4.2 E lysate) only comprising of HC with no LC in the reducing dimension, and again with the molecular weight from the non-reducing dimension it

can be suggested that this is a HC dimer. The first and third bands within the lysate denoted by a red and blue asterisk (*) are seen in the medium sample in both dimensions. This result confirms that IgG and HC dimers are both secreted. The lower molecular weight band which consists of LC and HC in the media denoted by a purple * is a half molecule of IgG (HCLC) which is again supported by its molecular weight in the non-reducing dimension. This result confirms that the protein band in the medium which was hypothesised to be a HC dimer consists solely of HC.

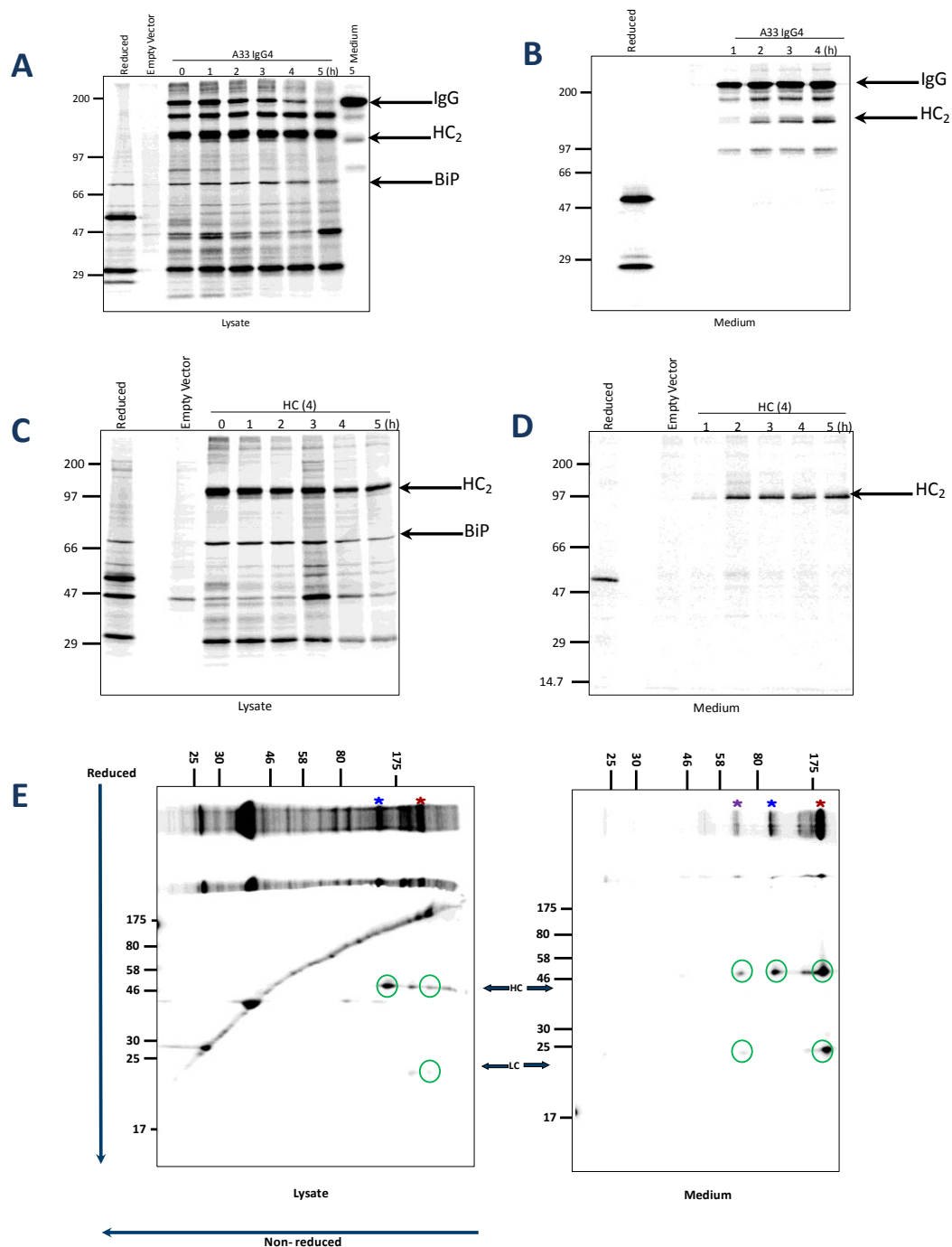


Figure 4.2: Assembly and secretion of A33 IgG4 analysing the composition of various antibody intermediates.

Pulse chase analysis of co-transfection of HC and LC (A+B) or HC alone (C+D) shows the secretion of HC dimers in the presence and absence of LC over a chase time of 5h. BiP can be seen in both transfections in the lysate as it co-precipitates with the unfolded HC. A chase time of 3h of a co-transfection with LC and HC underwent 2D gel electrophoresis to determine the composition of the protein bands (E). IgG, HC₂ and a half molecule of IgG can be identified denoted by the green circles in panel E. It can be seen that the protein band which was hypothesised to be a HC dimer solely contains HC in the medium.

As described earlier in the introduction, the folding of the HC is governed by the CH1 domain, which is said to require the CL domain of the LC to do so. The initial investigations of the HC dimer secretion suggests that the CH1 domain is folding in the absence of LC due to the faster migration of the protein in the medium when compared to the lysate of the cell. In order to access the folding status of the CH1 domain a point mutation was introduced to inhibit the domain from being able to fold. It has previously been shown that when the conserved proline residue which is isomerised during the folding process is mutated to alanine, this inhibits the ability of the CH1 domain to fold (Feige et al., 2009). Using this information and after studying the sequence of IgG4 CH1 domain, primers were designed to introduce the point mutation P151A to the CH1 domain of the A33 IgG4 HC. With this mutation successfully introduced we tested whether the mutant HC could be secreted. The HCs of both the WT and P151A mutant A33 IgG4 were separately transfected into CHO cells and incubated for 48 h. The cells were then radiolabelled with ³⁵S cysteine/methionine media for 30 minutes and chased for three hours with radiation-free media. The lysate and media were then affinity purified using protein A Sepharose beads and run under non-reducing conditions on a gradient (4-20%) SDS page gel (Figure 4.3 A). The A33 IgG4 HC WT showed secreted HC dimers as seen previously; however there was no evidence of HC dimers in the medium sample of the P151A mutant (the cleavage product described in Chapter 3 is also present with this mutation). This confirms the hypothesis that inhibiting the folding of the CH1 domain does inhibit the secretion of HC dimers and, therefore, confirms that the CH1 domain is folded in secreted HC dimers in the absence of LC.

However, as these are cultured cells, one could argue that the cause of the presence of HC dimers in the medium samples is actually due to cell lysis and are not in fact being actively secreted from the cell. In order to address this possibility glycosylation studies were carried out on the lysate and media samples of A33 IgG4 HC. As a protein travels through the secretory pathway different glycosylation events take place. In the ER of the cell proteins are glycosylated via an asparagine residue with mannose, glucose and N-acetyl glucosamine which are added as a pre-formed glycan core (Krambeck et al., 2009). The glucose molecules are trimmed as the protein goes through various quality control checkpoints within the ER where the correctly folded protein can then be transported to the Golgi apparatus (GA) and eventually secreted from the cell. However, within the GA the glycosylated protein can undergo more complex modifications before it is actively secreted. With this information in mind, two enzymes were used to test the glycosylation status of the HC dimer which contains a glycosylation site within the CH2 domain. The first was Endo H which is an enzyme that cleaves asparagine linked mannose rich oligosaccharides, typically like those found on proteins that exist in the ER. The second was PNGase F which is an enzyme that cleaves all N-glycans regardless of the extent of processing. Therefore Endo H will only be able to

cleave oligosaccharides on proteins which are present in the ER and not those that have passed through the GA (Freeze & Kranz, 2010).

CHO cells were transfected with IgG4 A33 HC and incubated for 48 h. The cells were then divided into three separate samples; the media was extracted and the cells were lysed. All six samples of lysate and media were affinity purified using protein A Sepharose to isolate the HCs which were then denatured and reduced. Two samples (one lysate, one medium) were incubated in buffer alone as a control, another two were incubated with Endo H and the last two with PNGase F. The samples were then run on a reducing gradient (4-20%) gel and silver stained (Figure 4.3 B). When comparing the samples within the lysate of the cell, the two protein bands corresponding to HC have a faster mobility through the gel after treatment with Endo H or PNGase F when compared to the buffer alone control. This result indicates that the HC is within the ER of the cell as it is cleaved by Endo H. Conversely, the samples of the medium show that upon the addition of Endo H the HC protein band migrates to the same height as the buffer alone control. However, upon treatment with PNGase F it can be seen that the HC has a faster mobility through the gel as its glycan has been cleaved. Therefore, as Endo H was not able to cleave the HC in the medium, this result confirms that the HC dimers are being actively secreted from the cell as they have passed through the GA and are therefore not a product of cell lysis.

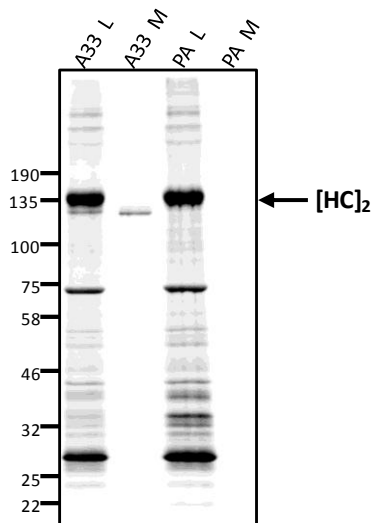
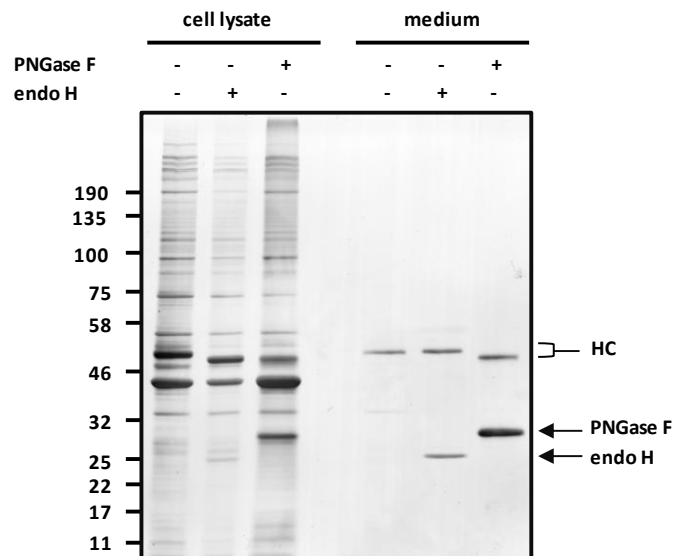
A**B**

Figure 4.3: Confirmation of HC dimer secretion via mutagenesis and glycosylation studies.

In order for the CH1 domain to fold, an isomerisation of a conserved proline residue must occur. This proline was mutated to alanine to prevent the CH1 domain from folding and is denoted PA. The WT HC with the variable region A33 and the PA mutant were separately transfected and underwent a pulse chase with a chase time of 3h. The lysate and medium samples were run side by side and analysed on a gradient (4-20%) polyacrylamide gel under non-reducing conditions. It can be seen that the WT secretes HC dimers and the PA mutant does not (A).

To confirm that the HC dimers are actively secreted and therefore are processed in the GA, glycosylation studies were carried out on a single transfection of A33 IgG4 HC. The lysate and medium were treated with either Endo H or PNGase F. Both enzymes cleaved the intracellular material, however only PNGase F cleaved the secreted material. The samples were analysed on a gradient (4-20%) polyacrylamide gel under reducing conditions and silver stained (B).

4.5 Sequence variation of the VH domain affects the level of secretion of the A33 HC dimer

HC dimer secretion was seen for both the isotopes IgG1 and IgG4 containing the variable region A33. This result deviates from the central dogma of antibody folding as the CH1 domain should only fold in the presence of the CL domain of the LC. The antibodies investigated are engineered and are, therefore, not naturally occurring; this may contribute to the secretion of HC dimers. The A33 IgG4 HC which has been studied is a chimeric HC consisting of a human sequence for the constant domains CH1, CH2 and CH3; however the sequence of the variable domain is murine. A chimeric antibody is produced as a result of humanisation to avoid an immune response if used as a therapeutic for humans.

The humanisation process of antibody engineering for therapeutic use involves several processes. Initially, an animal is inoculated with an antigen of choice to enable the production of antibodies with a strong affinity for the antigen. The splenic cells are then extracted and fused with myeloma cells to generate a hybridoma cell line which can be screened for the antibody (de StGroth & Scheidegger, 1980). Once the antibody sequence has been identified, the variable domain sequence combined with the human constant domains can be synthesised for use as a chimeric therapeutic. However, the variable domain will contain a foreign sequence which can cause an unwanted immune response. Therefore, in order to prevent this, the variable region can be further engineered in one of two ways; either the framework sequence of the variable domain is mutated to fit with the human sequence or the complementary determining regions (CDRs) in between the framework sequences are grafted onto the backbone of a human variable domain (Jones et al., 1986). Usually a combination of both is required to create a stable VH domain which does not cause an immune response, however the introduction of mutations can have a detrimental effect towards the function of binding to an antigen (Ewert et al., 2004). Either process is necessary to use the antibody as a therapeutic for treatment of humans in order to avoid an immune response against the antibody which would inhibit its therapeutic use. The A33 IgG4 HC investigated thus far has been humanised as a chimeric with no additional engineering so the sequence of the variable domain represents that of a mouse. Therefore, in order to investigate any consequences of furthering the humanisation process of this antibody a humanised version of A33 IgG4 HC was generated with a variable region which is more representative as human. For the purposes of this investigation the original A33 IgG4 HC is denoted as A33c for chimeric and the HC which has additional human engineering is denoted as A33h for humanised.

The DNA sequence alignment between the two variable domains illustrated 11 amino acid changes outside of the CDRs: K3Q, K13Q, K19R, T40A, E42G, R44G, A75S, R76K, S84N, S88A and L93V (Figure 4.4 A). These changes will have made the mouse backbone read more as human but will not affect antigen binding. CHO cells were transfected with A33c (chimeric) and A33h (humanised) IgG4 HCs separately and incubated for 48 h. The cells were then separated from the medium and the HCs were isolated using protein A Sepharose beads. Each sample was denatured and run under non-reducing conditions on a 10% SDS PAGE gel and developed via western blot using a fluorescently labelled protein A (Figure 4.4 B). When comparing the level of LC independent HC dimer secretion between the humanised HC and the chimeric HC it can be seen that the amount of protein secreted is greatly reduced for the humanised HC. Therefore, the difference of 11 amino acids within the variable domain can greatly reduce the amount of HC dimers secreted in the cell. This result suggests that the variable domain of the HC plays a role in the mechanism behind the secretion of HC dimers.

A



B

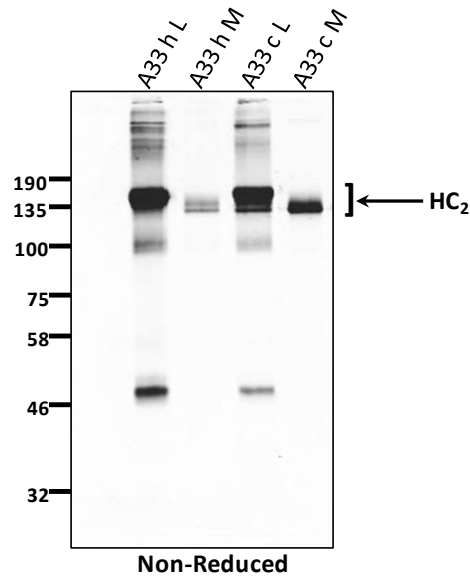


Figure 4.4: Contrast of variable region engineering via the humanisation process.

The sequence alignment of the original VH domain from the mouse against the same VH domain after humanisation yielding 11 amino acid differences (A). Western blot analysis of two A33 HCs one humanised and the other a mouse human chimeric looking at intracellular (L) and secreted (M) HC dimers (B) The same number of cells were affinity purified using protein A Sepharose beads prior to loading equal amounts on the polyacrylamide gel.

4.6 HC dimer secretion is variable region dependent

As described in the previous section, changes within the variable domain of the HC can be seen to reduce the amount of HC dimers secreted. To investigate this further, another six HCs were investigated (generation of HCs described in Chapter 3). All seven HCs have identical constant domains which are human and are of the isotype IgG4; CH1, CH2 and CH3. However, the variable domains all differ in their sequence and species originality (see table 1 for details). Therefore, any differences in the secretion of HC dimers can be attributed to differences between the VH domains.

CHO cells were separately transfected with each of the seven HCs plus the P151A A33 mutant to act as a negative control. The P151A mutant as mentioned in the previous section has been shown to lack the ability to secrete HC dimers due to the inability of its CH1 domain to fold. After a 48 h incubation period, the cells and medium were separated and any HCs within each sample were isolated by affinity purification using protein A Sepharose beads. Each sample underwent gel electrophoresis on a 7.5% SDS Page gel under non-reducing conditions with the media sample of each HC run next to its cell lysate (Figure 4.5 A). Of the eight HCs investigated, four are seen to secrete HC dimers; A33, 497, HC3 and HC4 and four do not; HC1, HC2, HC5 and P151A mutant negative control. All eight HCs had similar levels of protein expression as seen in the lysate samples. This is an intriguing result which strongly supports the hypothesis that the effect of HC dimer secretion is dependent on the variable region. Also, this result suggests variation in the level of HC dimer secretion between the four HCs when compared to the equal expression levels in the lysate. To confirm any variations in the amount of HC dimers secreted, the HCs were co-transfected with another protein which is known to be secreted; α 1 antitrypsin (α 1AT). A1AT is a 58 kDa protein containing no disulphide bonds. It does, however, contain several glycosylation sites. The varying levels of glycosylation within α 1AT cause multiple protein bands when developed on a western blot. The α 1AT construct used in these experiments has been modified with a V5 epitope tag for detection by western blot. By co-transfecting the HCs with α 1AT as a control, the relative amount of secretion of HC dimers between the four secretors can be quantified.

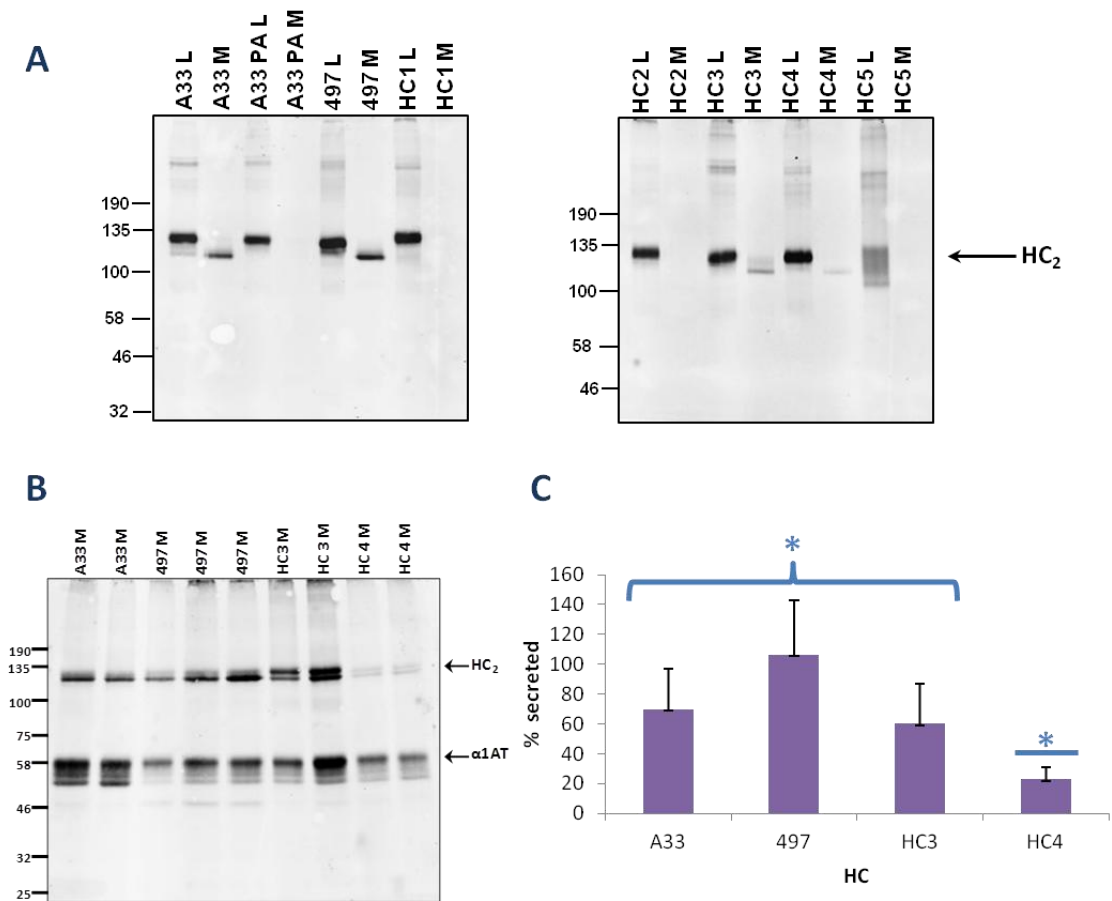


Figure 4.5: HC dimer secretion of seven different HCs of IgG4.

Seven HCs that differ in their variable regions plus the PA mutant of A33 which inhibits HC dimer secretion were transfected into CHO cells. The HCs were isolated from the medium (M) and lysate (L) using protein A Sepharose and run under non reducing condition with the corresponding medium sample run next to the lysate (A). Of the seven HCs A33, 497, HC3 and HC4 were seen to secrete HC dimers.

The four HCs which could secrete HC dimers were co-transfected with $\alpha 1AT$ and TCA precipitated from the medium. The amount of each HC secreted was compared to $\alpha 1AT$ as a loading control (B) and quantified (C). As each cell should take up the same amount of DNA (HC + $\alpha 1AT$) and the level of expression between different transfections of $\alpha 1AT$ should be equal; the level of relative secretion can be compared to the amount of $\alpha 1AT$ as a loading control. The quantifications were carried out using $n=5$ and a T-test was used to measure the significance. The T-test used was type 1 with 2 tails and measured by dividing one data set by the other. The quantifications showed that A33 (T-test = 0.03), 497 (T-test = 0.004) and HC3 (T-test = 0.04) all significantly secrete HC dimers to a greater extent than HC4 (C).

CHO cells were co-transfected with each HC plus α 1AT and incubated for 48 h. The medium was then separated from the cells and the proteins were extracted using TCA precipitation (see Chapter 2, section 2.3.5). Each sample was run on a 12.5% polyacrylamide gel and developed using an anti-V5 antibody to detect α 1AT and a fluorescent protein A to detect the HC. This experiment was repeated five times to acquire a statistically viable result. Each protein band for HC present in the medium was quantified and normalised against α 1AT (Figure 4.5 B). Differences in secretion between A33, 497 and HC3 were inconsistent so placing them in order of most secreted to least was not possible. However, A33, 497 and HC3 all consistently secrete HC dimers at a higher level than HC4 (Figure 4.5 C). This result expands on the concept that not only is HC dimer secretion dependent on the variable domain, but the level of secretion is also varied between the HCs which do secrete. To understand what these differences are the sequences of the VH domains were studied and compared (Figure 4.6 A). The variations between all seven HCs are so extensive that to narrow them down the two were separated into “secretors” and “non-secretors”. Comparison of the sequences shows differences between them, however, one consistency when comparing the secretors to the non-secretors is the length of the CDRs; particularly CDR3.

CDRs of the variable domain of an antibody are the regions which participate in antigen binding. In the case of the HC secretors the CDR3 is shorter than those of the non-secretors. This difference could add to the proteins ability to fold the CH1 domain and secrete HC dimers. The CDRs of the VH domain are comprised of the sequence of three separate genes in antibody development; V (variable), D (diversity) and J (joining). Three random genes from each repertoire (of which there are approximately 51 potential V sequences, 27 potential D sequences and 6 potential J sequences) are selected and joined together to create the variable domain; the V gene encoding for CDR1 and 2 and all three encoding CDR3 many of which have been sequenced and are available as a library for identification (Knappik, Ge et al., 2000). The variable domain then goes through several more stages of mutation to yield an antibody which tightly binds to its target antigen. As the CDR3 is comprised of V, D and J genes it is not surprising that this region of the HC has the most variation. The length of CDR3 is attributed to the amount of junctional flexibility the variable domain undergoes which could explain the differences in lengths we see between the secretors and non-secretors. The specific V D and J genes which comprise each HC can be identified via sequence analysis and using the online antibody sequence database IMGT (Figure 4.6 B). Upon initial inspection of the various V D and J combinations there does not seem to be an obvious trend between the HCs which secrete dimers, however all three non-secretors either have DH1 or JH2 as their recognised D or J gene. As this is a comparison on a minuscule scale compared to the 10^{10} possibilities of potential variable domains generated then one cannot be

conclusive of a trend between only three. However, within the pool of HCs investigated there does seem to be some trends which could suggest specific gene families which are contributing or not to LC independent HC dimer secretion.

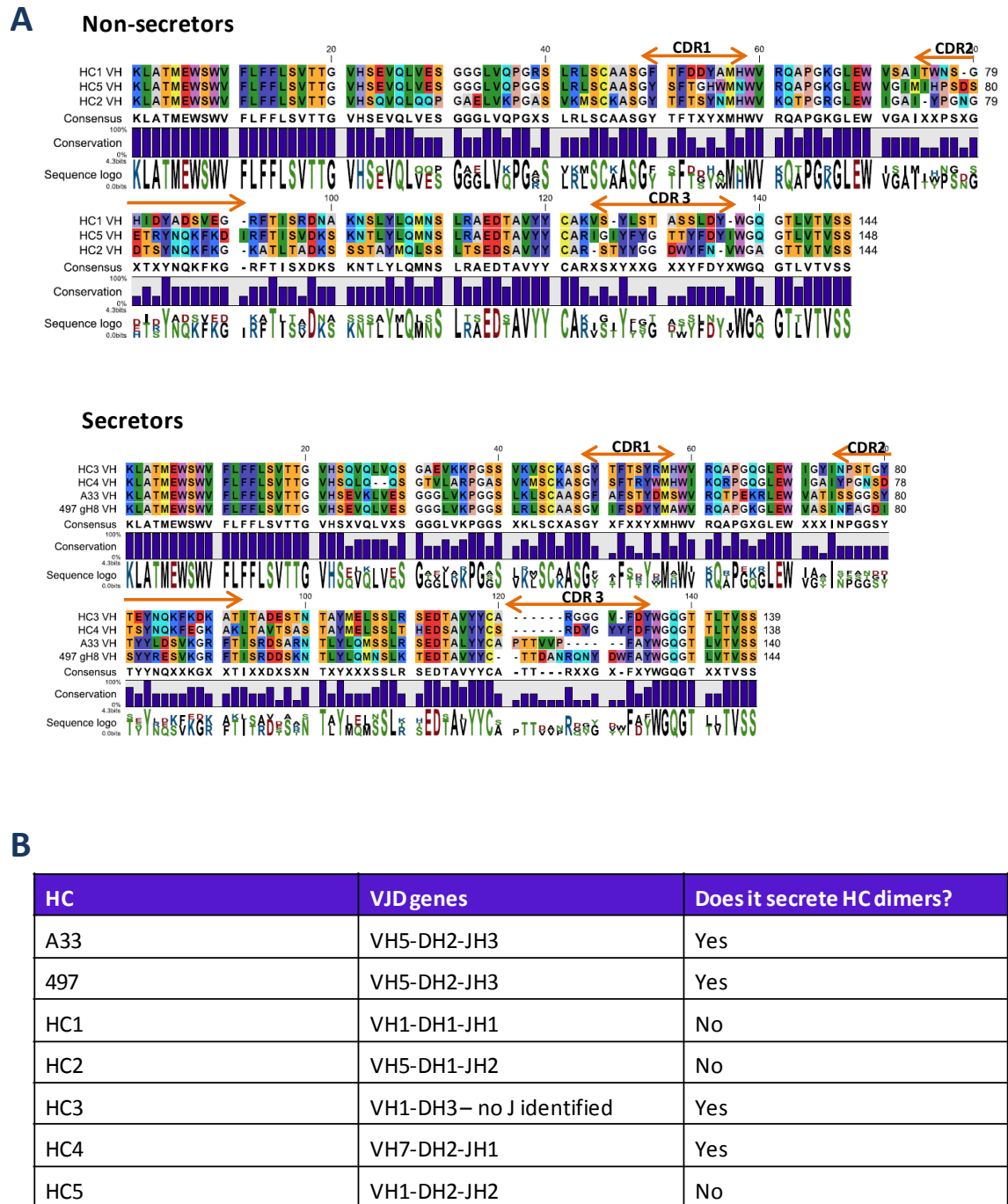


Figure 4.6: Sequence alignments and VDJ assignments of the seven VH domains investigated.

Amino acid sequence alignments of non-secretors and secretors highlighting the differences between them with the CDR regions highlighted (A). It can be seen that the length of the CDR3 is shorter for the secretors when compared to the non-secretors. VDJ assignments of each VH domain investigated also showed a slight trend; all non-secretors contain either DH1 or JH2 (B).

4.7 LC dimer secretion is variable region dependent

As well as looking at the HCs ability to fold and secrete dimers, two LCs were also investigated for comparison. It has been previously shown in the literature that some LCs have the ability to fold and secrete disulphide linked dimers via interactions between the CL domains (Roussel, Spinelli et al., 1999). In order to determine if this effect is seen with any of the IgG constructs investigated, two LCs from known HC dimer secretors were examined; A33 and 497. Both LCs have identical human κ CL domains but differ in their VL domains.

CHO cells were transfected with A33 LC or 497 LC and incubated for 48 h. The cells were then radiolabelled with ^{35}S cysteine/methionine media for 30 minutes and chased with radiolabel free media for 3 h. The lysate and media of each sample was collected and immunoisolated using protein A Sepharose beads incubated with anti κ LC. Each sample was then denatured, reduced and underwent gel electrophoresis on a gradient (4-20%) polyacrylamide gel (Figure 4.7 A + B). The LCs of both A33 and 497 were successfully expressed in the cell lysate, however, there is a huge reduction (95%) of LC present in the medium of A33 compared to 497 (Figure 4.7 C). The partner HCs of the two LCs both secrete similar levels of HC dimers in the absence of LC, conversely in the absence of HC, the ability of the LCs to be secreted is different. It was also demonstrated that the LC secreted by 497 exists as a dimer when the sample is analysed under non-reducing conditions (Figure 4.7 D). This result highlights the fact that antibody folding is much more complex than originally thought with each variable domain of both the LC and the HC potentially controlling the secretion of incomplete IgG in the form of LC or HC dimers.

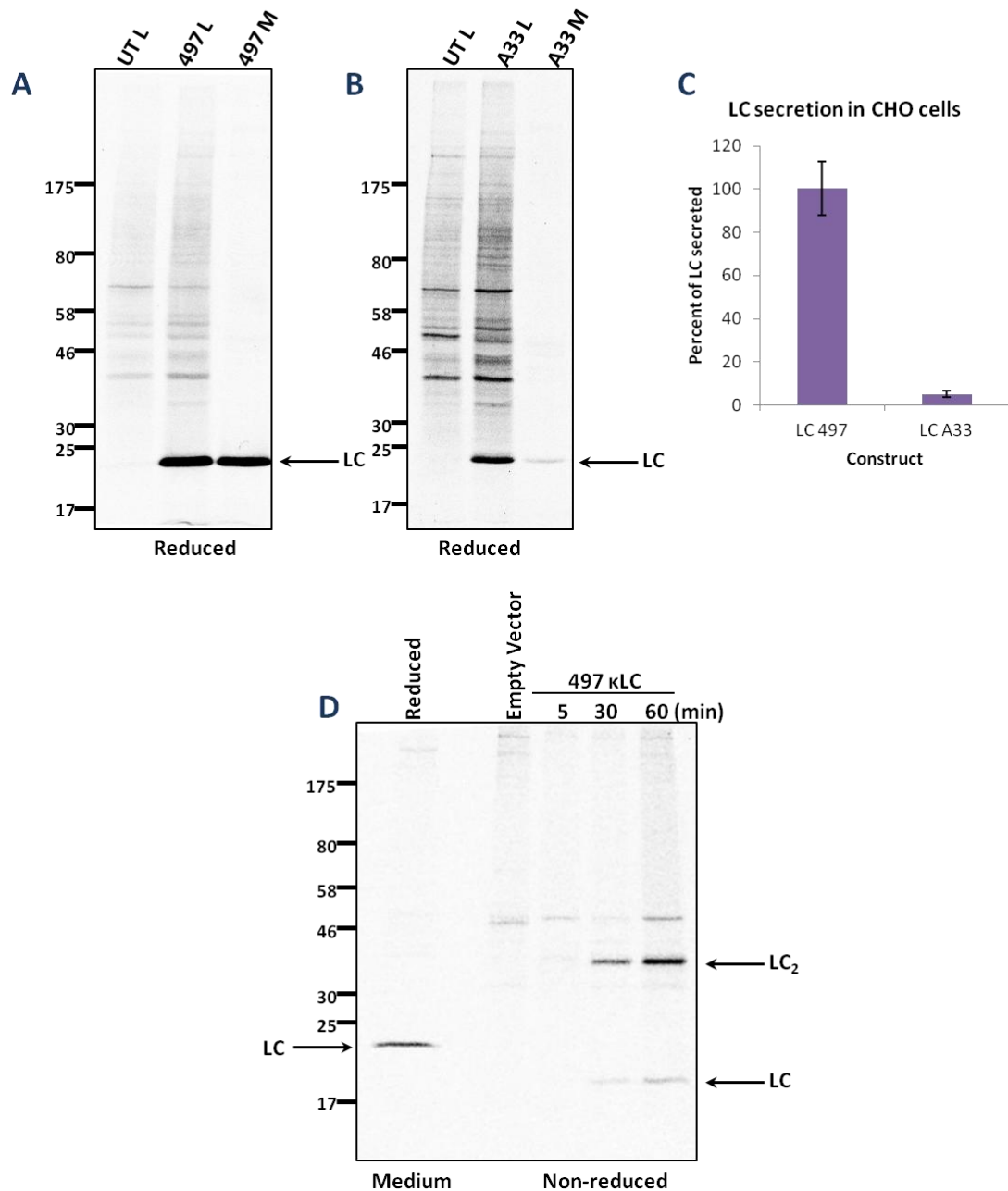


Figure 4.7: LC assembly and secretion of 497 and A33 in the absence of HC.

The LC of 497 (A) and A33 (B) underwent pulse chase analysis for a chase time of 3h and the intracellular (L) and secreted material (M) were analysed under reducing conditions. It can be seen that the LC of 497 secretes 100% LC whereas A33 secretes considerably less protein at only 5% (C). Quantification of the percentage of secretion was achieved by comparing the secreted material to the intracellular material using an n=3 and a T-test for statistical significance with a value of 0.000178. Observing LC secretion for 497 under non-reducing conditions shows that the LC is secreted as a disulphide linked dimer (D).

4.8 Generation and Secretion of Fab Fragment of A33 IgG4 HC

Previous work described in this chapter has shown that the HCs of IgG can form dimers and are actively secreted from the cell in the absence of LC. This work has also shown that the secreted HC dimers have a folded CH1 domain. Therefore, the next question to address is whether or not the Fc fragment of the HC, which is known to associate and stabilise the HC dimer, is required for the process of HC dimer secretion.

A construct coding for the CH1 domain of IgG4 and the variable domain A33 was generated with a V5 epitope tag for detection and isolation (Figure 4.8 A). The cDNA sequence of the full length HC of A33 IgG4 was studied and primers were designed to create a PCR product with different restriction sites of HindIII and NotI plus a V5 tag C terminal to the Fab sequence. The coding sequence of the Fab fragment ends at the CH1 domain, and contains no hinge sequence. This coding sequence was then sub-cloned into a pcDNA 3.1 vector for expression. CHO-S cells were transfected with this Fab HC and the medium and cell lysate were collected. The protein was immunoprecipitated from both fractions using anti-V5 agarose beads. These samples were then denatured and separated under both reducing and non-reducing conditions (Figure 4.8 B).

Secretion of the truncated HC can be seen which demonstrates the Fc fragment is not a prerequisite for HC dimer secretion. The result shows that under non-reducing conditions the Fab HC mainly exists as a monomer within the cell lysate migrating at approximately 23 kDa; however, upon secretion the HC exists solely as a dimer (46 kDa). This dimer is not present when the proteins are run under reducing conditions indicating that the Fab dimer contains a disulphide bond. This result illustrates that the Fc fragment of the HC is not required for the formation or secretion of HC dimers. In order to determine which cysteine is forming the disulphide and to address the question of whether secretion is possible without the presence of a disulphide bond a C131S mutant was created within the CH1 domain. This mutant abolishes the protein's ability to form that disulphide but as the two amino acids are chemically similar the protein should still be able to fold and retain its function. The C131S mutant was transfected and isolated in the same manner as the wild type (WT) (Figure 4.8 B last two lanes of each gel). The proteins were again examined under both reducing and non-reducing conditions; in this case the protein only existed as a monomer regardless of the presence of a reducing agent. Secretion of the HC also occurred for the mutant, however although the interchain disulphide is not present the HCs could still potentially exist as a dimer through non-covalent interactions. This result demonstrates that the disulphide bond is not required for secretion of the Fab HC as the protein is secreted as a monomer as it can no longer form a dimer without cysteine 131. This implies that the CH1

domains are associating with each other in the form of a dimer and that this association does not require the formation of a disulphide.

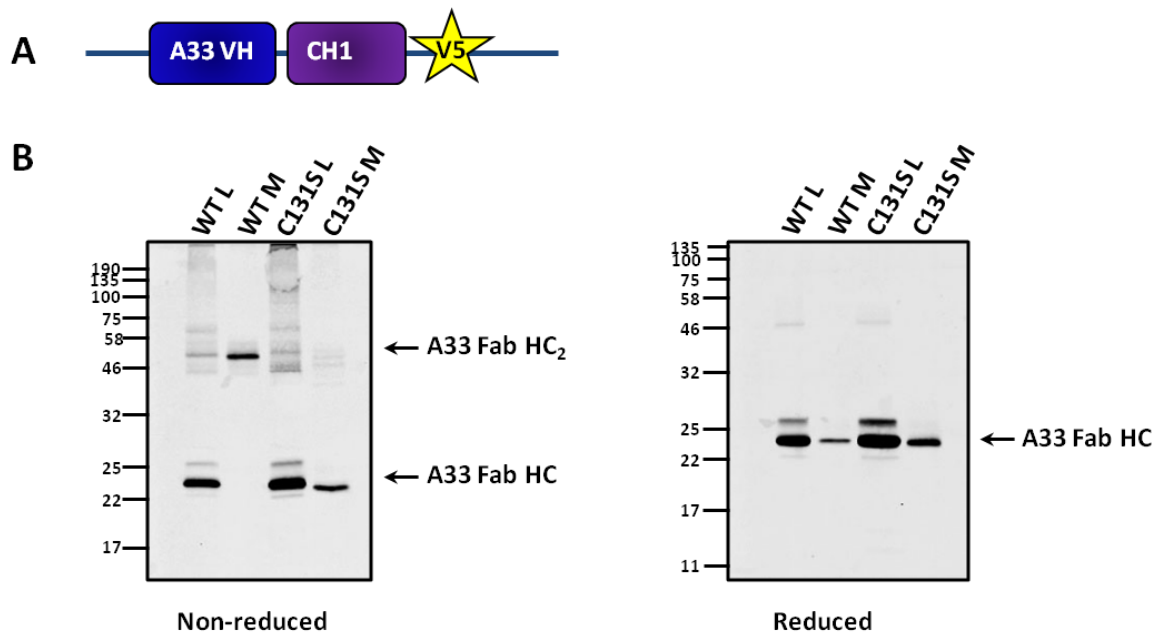


Figure 4.8: Generation and evaluation of the assembly and secretion of the truncated A33 IgG4 HC.

A truncated antibody Fab fragment (A33 VH domain and CH1 domain) with a V5 tag was generated from A33 IgG4 HC (A). This construct plus a C131S mutant which is unable to form a disulphide linked dimer were transfected into CHOS cells and run under non-reducing and reducing conditions via western blot using an anti-V5 primary antibody for detection (B).

4.9 The A33 variable domain folds the CH1 domain by means of an alternative mechanism which is used by the CL domain

The results presented in this chapter have revealed that HC dimers are actively secreted in a LC independent manner. It has also been described that this process is variable region dependent implicating this domain as a factor in the mechanism behind HC dimer secretion. On the basis of the evidence presented thus far, two potential mechanisms describing how the variable domain aids in the folding of the CH1 domain can be hypothesised. The first is that the Fab arms are interacting in *trans* of one another meaning the VH domain of the opposite Fab domain is acting as a template to fold the CH1 domain in a similar manner as to how the CL domain is able to act in the folding of the CH1 domain. The second potential mechanism is that the opposite variable domains are interacting with each other enabling the two CH1 domains to become close enough to stabilise each other to enable folding to occur (Figure 4.9). Ultimately the mechanism behind the folding of the CH1 domain could theoretically be a combination of the two with the VH domain introducing some structure to the CH1 domain on the same polypeptide, stabilising it enough to interact with the other CH1 domain to enable disulphide formation and complete folding to occur.

In order to test this hypothesis, the isolated domains of immunoglobulin were studied. This work originates from a previous study which involved purification of a CH1 domain from IgG1 and a CL domain, which were mixed together to demonstrate that the CL domain can fold the CH1 domain. Folding was assayed by studying the circular dichroism (CD) of the proteins either in isolation or as a complex (Feige et al., 2009). To expand on this work, the experiment was repeated using the human CH1 domain from IgG4 and the human kappa CL domain which have been used in all experiments that have been undertaken during the investigation of HC dimer secretion. Also, to investigate the folding of the CH1 domain in the absence of LC the murine variable domain of A33 was also purified to incubate with the CH1 domain in the same manner as the CL domain to test whether it has the ability to introduce structure to the protein.

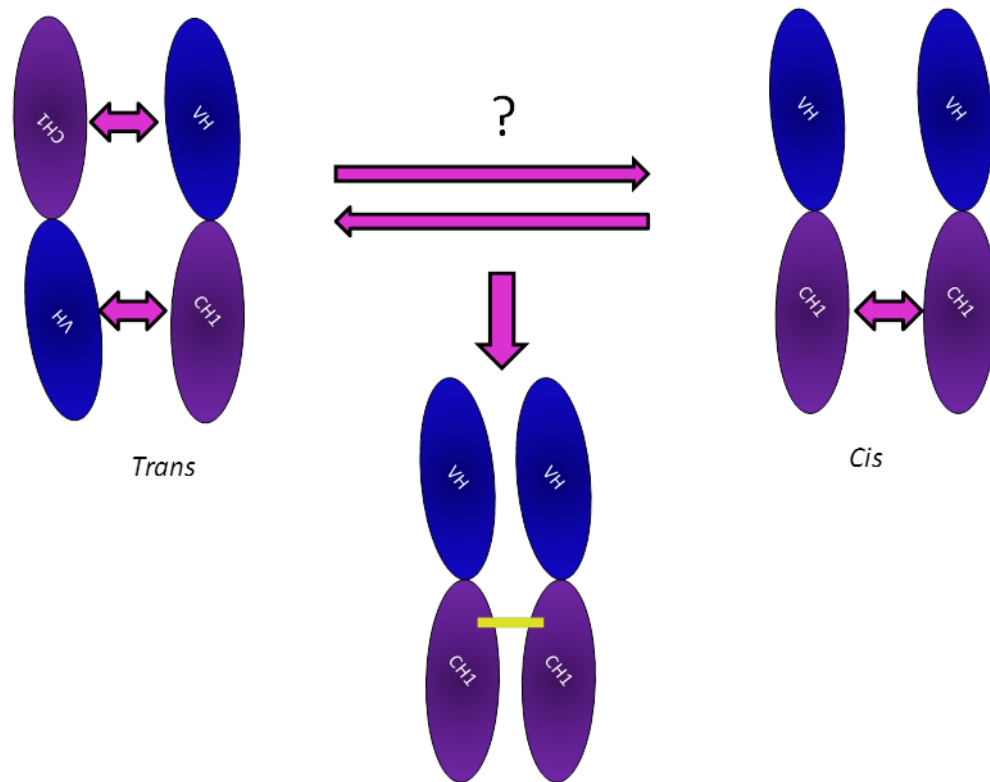


Figure 4.9: Two potential mechanisms as to how the VH domain is folding the CH1 domain.

The Fab HCs could either be interacting in *trans* of one another, for example the VH domain acting on the opposite CH1 domain in the same manner as the CL domain to introduce structure into the protein. The alternative scenario is that the two VH domains are interacting with one another, enabling the CH1 domains to come together in *cis*. The VH domains are able to stabilise the CH1 domains enabling them to fold and form an interchain disulphide.

4.9.1 Purification of individual immunoglobulin domains

First, the three domains were codon optimised for expression in *E.coli* and cloned into a pET28a expression vector which allows for expression of an N-terminal His₆ tagged recombinant protein. Initially the solubility of each protein was tested by expression in BL21 Star (DE3) pLysS strain of *E. coli* and inducing expression with 0.1 mM isopropylthio- β -galactoside (IPTG) at 16°C once the cultures had reached a mid-log phase of growth. IPTG is a lactose homologue which is used in expression systems in *E.coli* as it activates the lac operon by binding to the lac repressor, thereby controlling when the protein of choice is expressed (Kaern, Blake et al., 2003). The initial cell lysis was carried out by resuspending the cell pellet in lysis buffer to generate a whole cell lysate (WCL). After complete lysis using the freeze-thaw protocol described in materials and methods a solubilised fraction (SF) was collected (Figure 4.10). Each sample was reduced, denatured and underwent gel electrophoresis on a 15% polyacrylamide gel. The proteins were detected using an anti-His antibody via western blot analysis. It can be seen from the gel that the majority of the A33 and CL protein exists in the SF and can therefore be purified as soluble protein, however there is very little CH1 that is soluble and will be expressed as inclusion bodies. This is not surprising as the CH1 domain is known to be disordered and expressed as inclusion bodies (Feige et al., 2009). As the CH1 domain is expressed as inclusion bodies the method used to express the protein was changed to 1 mM IPTG at 37°C in order to optimise the amount of protein produced by the *E.coli*. The inclusion bodies containing the CH1 domain were prepared in a guanidine hydrochloride (GdHCl) buffer and loaded onto a His High Trap 5 ml nickel agarose column (processing of inclusion bodies and full protocol Chapter 2.42). The protein was refolded on the column by slowly removing the GdHCl. The protein was then eluted using an imidazole gradient. The elution profile from the CH1 domain shows a gradual increase in absorbance over time with an increasing concentration of imidazole, however, it does not yield a specific peak (Figure 4.11 A). In order to observe where the protein has eluted, the fractions were separated on a 15% polyacrylamide gel under reducing conditions and stained with Coomassie blue (Figure 4.11 B). There was only one protein band visible in each fraction which migrated to a size of approximately 13 kDa which corresponds to the size of the CH1 domain. The possible cause for a range in the elution displayed by the CH1 domain could be as it is disordered, it may conform to different folding structures and intermediates. The fractions which contained the CH1 domain were pooled, concentrated and underwent buffer exchange using a bench top PD10 column (50 mM Tris 50 mM NaF 0.5 mM EDTA pH7.45) in order to prepare the protein for circular dichroism (CD).

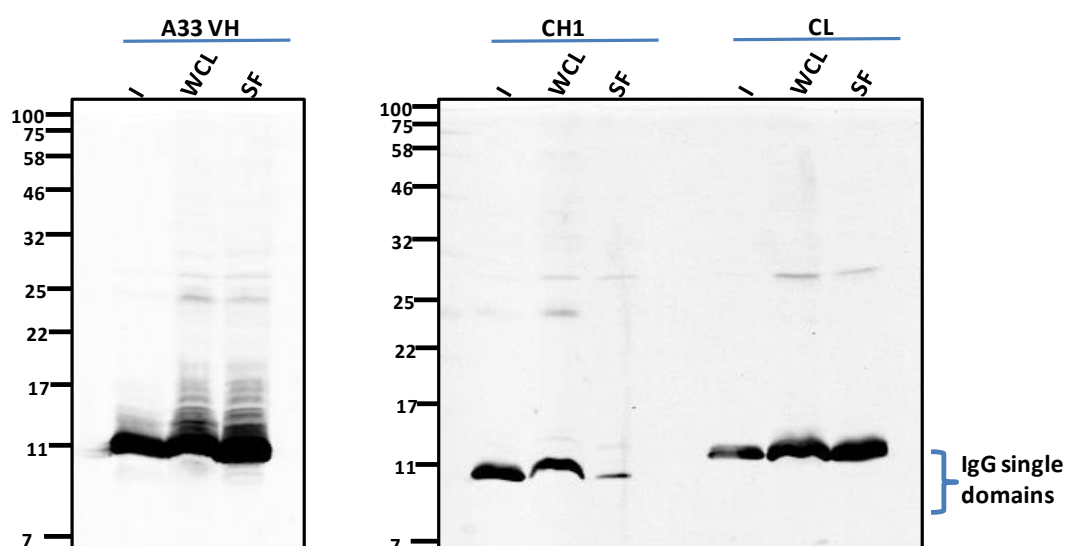


Figure 4.10: Solubility test of each individual immunoglobulin domain expressed.

Anti His western blot analysis of the immunoglobulin domains A33, CH1 and CL expressed in *E.coli*. The three lanes in each sample correspond to I = induced expression culture, WCL = whole cell lysate and SF = soluble fraction. The result shows that both A33 and CL domains express well in the soluble fraction, however the CH1 domain does not and therefore exists as inclusion bodies.

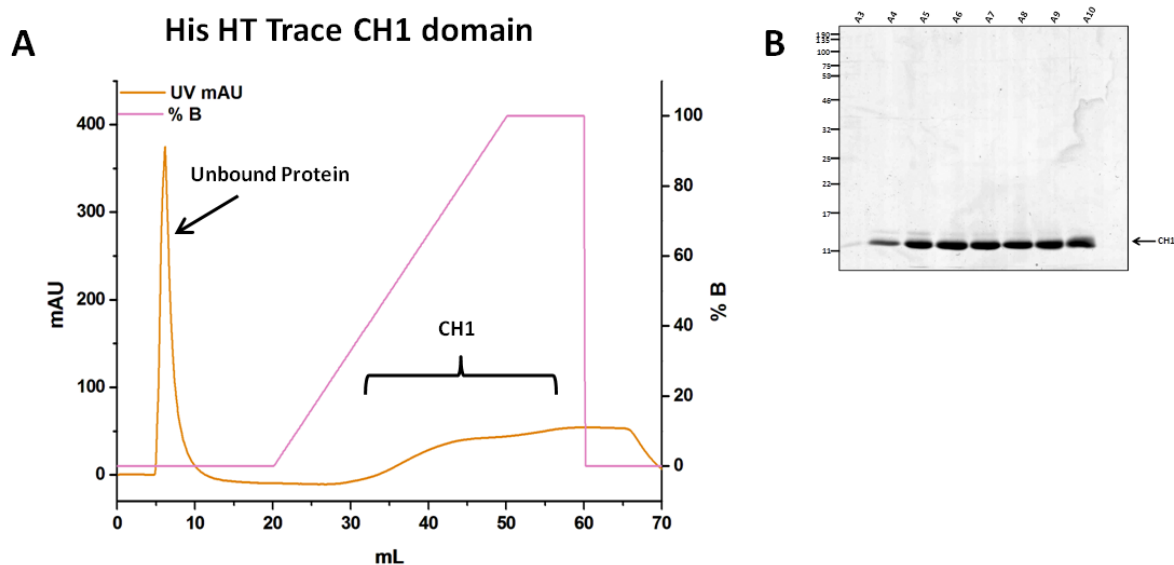


Figure 4.11: Purification of CH1 domain from inclusion bodies.

CH1 domain was purified from inclusion bodies on a His HT 5ml nickel agarose column where the orange line represents the UV mAU and elution profile of the protein and the pink line represents the linear gradient in % of the elution buffer B used (A). The fractions highlighted by the bracket in A were analysed via coomassie stain on a reducing 15% polyacrylamide gel (B).

The CL and A33 variable domain were both expressed and purified in a similar manner to each other. Each protein was soluble and expressed as described earlier. The soluble cell lysate of each domain was separately loaded onto a His HT 5 ml nickel agarose column and were eluted with an imidazole gradient. The elution profile for A33 from this initial purification yielded a single peak (Figure 4.12 A). The unbound protein in the flow through and wash steps, and eluted samples surrounding and including the peak were collected and separated on a 15% polyacrylamide gel under reducing conditions using Coomassie blue stain (Figure 4.12 B). The most abundant protein expressed within the fractions of the peak corresponded to the molecular weight of 15kDa. It can be seen that there are also higher molecular weight proteins which are contaminants that have eluted with the A33 VH domain.

Fraction A6 was used for the second purification step of gel filtration which separates proteins according to size and therefore would remove the higher molecular weight contaminants from the fraction containing A33 VH domain. The gel filtration column was calibrated previously using several protein standards and showed an optimal range to purify the size of the immunoglobulin domains (see appendix 4). The elution profile of A33 VH domain again yielded a single peak with an elution volume (V_e) of 17.696ml (Figure 4.12 C). This V_e corresponds to the molecular weight of the A33 VH domain of 15.06kDa when compared to the protein standards which were used to calibrate the column (Figure 4.12 D). To confirm the A33 VH domain was present in the peak the fractions were again separated on a 15% polyacrylamide gel under reducing conditions and stained with Coomassie blue (Figure 4.12 E). The results from the gel show the purification was successful as fractions A11 and A12 contain the A33 VH domain and the higher molecular weight contaminant proteins had been removed.

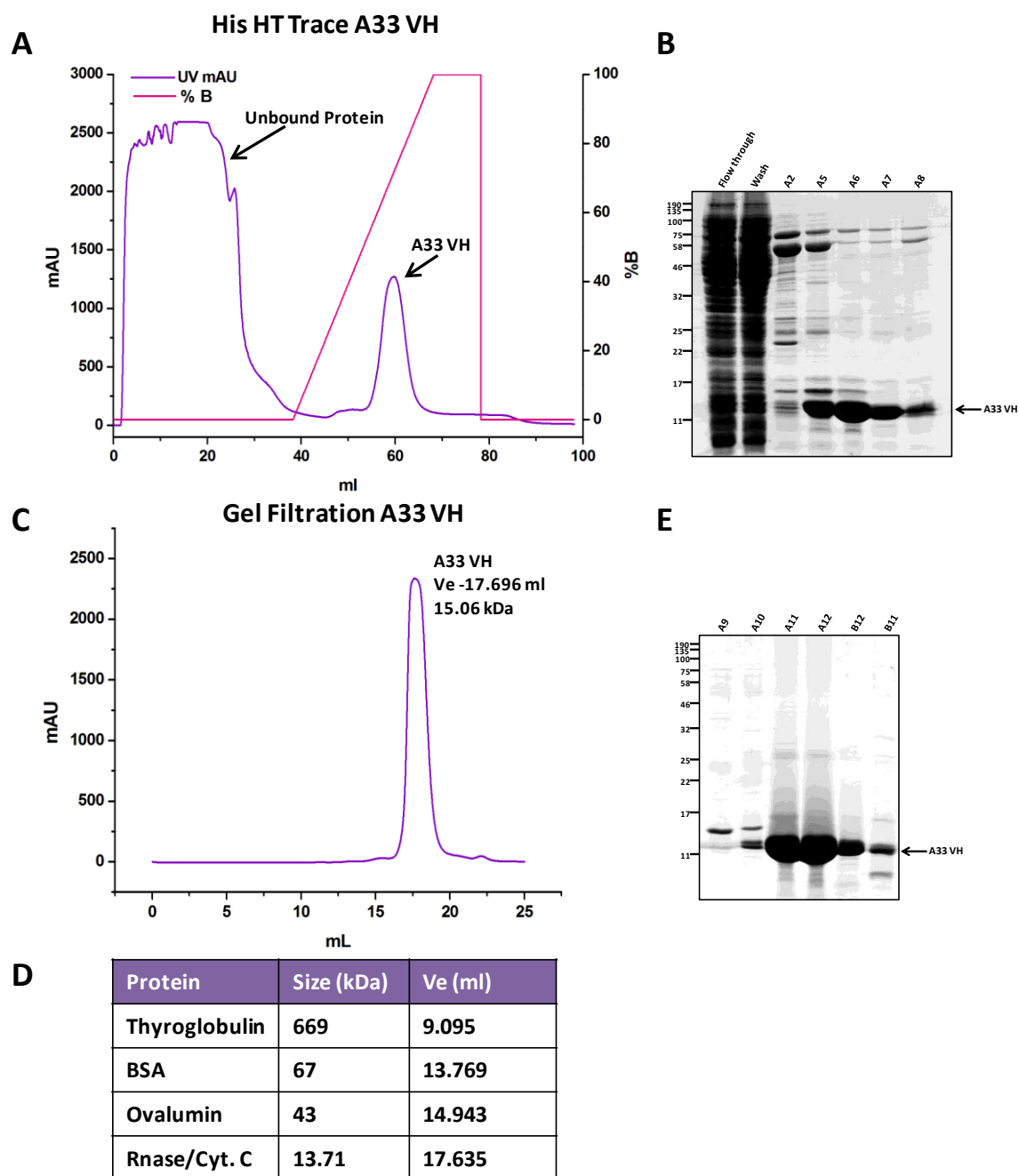


Figure 4.12: Purification and analysis of isolated A33 VH domain

A33 VH domain was purified as soluble protein on a His HT 5ml nickel agarose column, the purple line represents the UV mAU of the protein and the pink line represents the linear gradient of the elution buffer B as a percentage (A). The fractions were analysed via coomassie stain on a 15% polyacrylamide gel (B). Fraction A6 was then loaded onto a gel filtration (GF) column (C). Protein standards for the gel filtration column listed for comparison (D). GF was performed to purify the protein from the other contaminants and purity was confirmed via coomassie stain on a 15% polyacrylamide gel (E).

The CL domain was purified in the same way by firstly loading the soluble protein from the cell lysate onto a His HT nickel agarose 5 ml column. The elution profile of the CL domain yielded two peaks (Figure 4.13 A). The fractions surrounding these peaks plus the unbound protein from the flow through and wash steps were collected and separated under reducing conditions on a 15% polyacrylamide gel stained with Coomassie blue (Figure 4.13 B). As seen with A33 VH, the CL domain is present in the second peak but contains several higher molecular weight protein impurities. To remove these, the samples containing the CL domain were pooled and underwent gel filtration. The elution profile yielded two peaks for the CL domain, the first with a V_e of 17.3 ml corresponds to the molecular weight of the CL domain of 13.6 kDa, the second likely being a lower molecular weight contaminant protein (Figure 4.13 C). The molecular weight of the CL domain was calculated to be 13.6 kDa after comparison to the standard curve (Figure 4.14 D). The presence of the CL domain within this peak was confirmed by Coomassie staining on a reducing 15% polyacrylamide gel where the majority of the higher molecular weight protein contaminants have successfully been removed (Figure 4.13 E). The sample was then used for CD analysis.

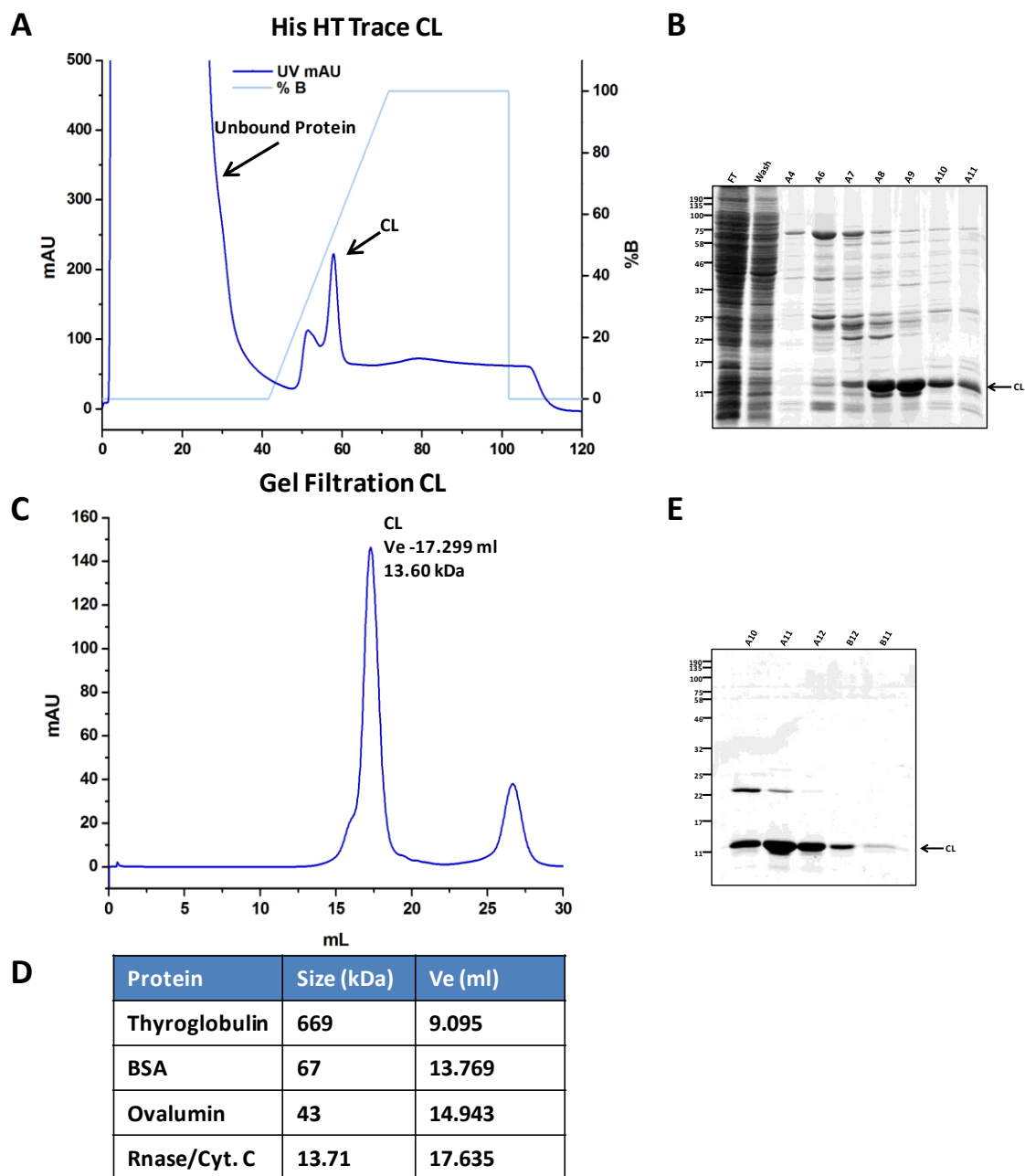


Figure 4.13: Purification and analysis of isolated CL domain.

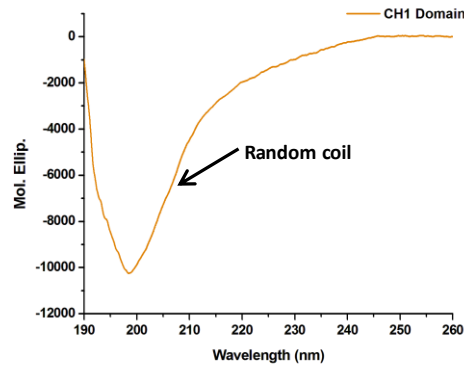
CL domain was purified as soluble protein on a His HT 5ml nickel agarose column (A). The fractions were analysed via coomassie stain on a 15% polyacrylamide gel (B). Fractions which contained the CL domain were pooled and then loaded onto a gel filtration column (C). Protein standards for the gel filtration column listed for comparison (D). GF was performed to purify the protein from the other contaminants and purity was confirmed via coomassie satin on a reducing 15% polyacrylamide gel (E).

4.9.2 CD of immunoglobulin domains in isolation

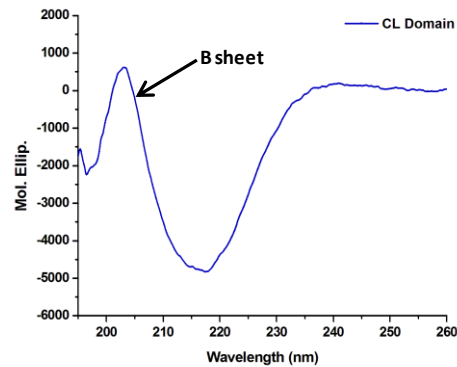
CD is a spectroscopic technique which measures the differential absorption of circularly polarised light by chiral molecules (Greenfield, 2006). CD has previously been used to study the structure of immunoglobulin domains in several studies which provides a comparison for our investigations (Kats, Richberg et al., 1995, Rousseaux, Aubert et al., 1982, Vermeer & Norde, 2000, Yageta, Lauer et al., 2015). The CD spectra of the CH1 domain of IgG1 and the CL domain have previously been shown by Feige and colleagues 2009, and therefore should be comparable to the results presented in this chapter.

The protein concentration of the CH1 domain was measured at 280 nm using a Cary 300 Bio UV-Visible spectrophotometer and was calculated to be 0.5 mg/ml using the extinction coefficient 11585. Far UV CD was measured using a Jasco J-810 spectropolarimeter and a quartz cuvette with a pathlength of 0.02 cm (Figure 4.14 A). The spectra were corrected for cell pathlength, protein concentration and buffer contribution. The spectrum shows a large trough at 200 nm with no positive maximum indicating the protein does not contain any ordered structure. These data compare well with spectra previously described by Feige and colleagues who saw the same spectra for the CH1 domain they purified.

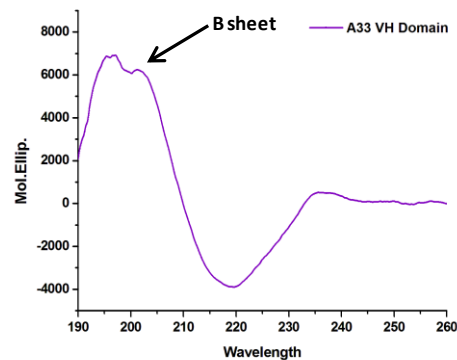
Far UV CD was performed on the CL domain and the A33 VH domain (Figure 4.14 B+C). Both spectra have a similar profile comparable to what has been previously described by Feige et al for the CL domain. There is a slight peak at approximately 230 nm which is more pronounced for A33, which may be due to aromatic amino acid contributions in this region. The most predominant peak for both proteins is between 200-210 nm which is indicative of β sheet suggesting that the secondary structure of both domains is mainly β sheet. This result demonstrates that there is folded protein in the protein purified from *E.coli*.

A

Far UV CD spectrum of CH1 (0.77 mg/ml; 0.02cm pathlength quartz cuvette).

B

Far UV CD spectrum of CL (0.71 mg/ml; 0.02cm pathlength quartz cuvette).

C

Far UV CD spectrum of A33 (0.32mg/ml; 0.02cm pathlength quartz cuvette).

Figure 4.14: Far UV CD Spectra of CH1 domain (A), CL domain (B) and A33 VH domain (C).

Far UV spectra of isolated domains taken from wavelengths of 190-260 nm. The CH1 domain spectra represents random coil indicating there is no ordered structure to the protein (A). The CL (B) and A33 VH (C) domains spectrum have a peak at approximately 200 nm which is indicative of β sheet.

4.9.3 Mixing individual CL and CH1 domains introduces structure into the CH1 domain which is not seen when the CH1 and A33 VH domains are combined.

The results presented thus far have shown the successful purification of the CH1, CL and A33 VH domains with the CH1 domain disordered and the other domains structured. Therefore, the hypothesis that the A33 VH domain is able to fold the CH1 domain in the same manner as the CL domain was examined. The CH1 domain was incubated with either the CL domain or the A33 VH domain for 4 h at 25°C. Both the CL and A33 VH domains were in excess of the CH1 domain at a ratio of 3:1. This ratio was used as these were the same conditions which were optimised by Feige and colleagues. In order to determine if there was any change in the CD spectra when the two proteins were mixed together, a theoretical sum of the spectra was calculated. The individual spectrum from each domain has been measured for the same incubation time, these could be added together to give a theoretical sum. Therefore, if any change is introduced to either protein upon incubation the spectrum would change from the original and would deviate from the theoretical sum. The theoretical sum was calculated for both mixtures and compared to the experimental spectra produced using far UV CD (Figure 4.15 A + C). The experimental spectrum of the CL and CH1 domain mix differs from the theoretical sum, indicating that a change has occurred in one or both of the proteins (Figure 4.15 A). Previous results shown by Feige et al indicated that the change occurred in the CH1 domain. In order to observe the contribution of the CH1 domain, the original CL domain data was subtracted from the experimental sum to leave only the spectrum for the CH1 domain (Figure 4.15 B). A difference can be seen between the isolated CH1 domains and when it is mixed with the CL domain as the spectrum has shifted to a more ordered state. However, this result is not as pronounced as seen with Feige et al. Upon comparison with the A33 however, there is no difference between the experimental sum and the theoretical sum of the A33 VH domain incubated with the CH1 domain (Figure 4.15 C). This result is the opposite of what was seen with the CL domain indicating that upon incubation with the A33 VH domain there is no change in structure to the CH1 domain. This was confirmed by taking the experimental spectrum and subtracting the single spectrum of the A33 VH domain to compare the CH1 domain prior and post incubation (Figure 4.15 D). There is no difference in the spectrum of the CH1 domain post incubation with the A33 domain. Therefore, as the A33 VH domain does not have the ability to introduce structure to the CH1 domain in the same way as the CL domain the mechanism behind HC dimer secretion is more likely to follow the *cis* pathway described earlier.

The A33 VH domain cannot fold the CH1 domain in *trans* like the CL domain yet it can fold the CH1 domain whilst on the same polypeptide chain. Therefore, the two variable domains must be interacting with each other in order to stabilise the CH1 domains to enable them to fold, as when the CH1 domain is incubated alone it still remains disordered. The proposed mechanism for the folding of the CH1 domain in the absence of LC is that the VH domain is stabilising the CH1 domain to allow folding. The two CH1 domains can then associate, dimerise and fold allowing secretion from the cell. This association may inhibit the ability of the LC to associate with the HC which is why the presence of HC dimers can be seen when HC is co-transfected with LC.

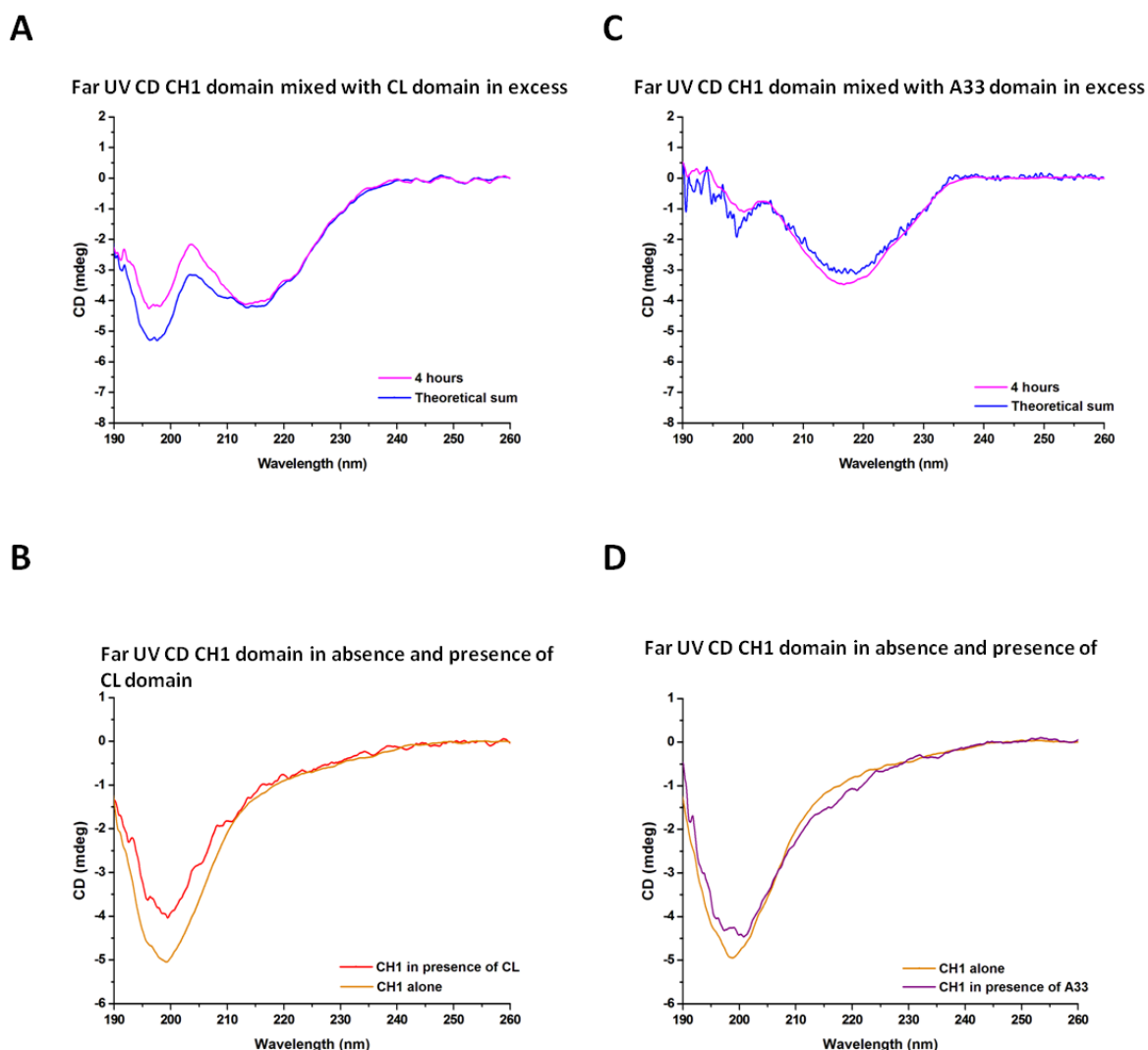


Figure 4.15: Far UV CD Spectra of CH1 domain mixed with either CL domain or A33 VH domain.

The theoretical sum of the spectrum of the individual immunoglobulin domains mixed together was calculated by adding the spectra previously calculated of the domains in isolation (blue line panels A + C). The experimental spectrum was measured after a 4h incubation of the domains mixed together at 25°C with the CL or VH domain in excess by a ratio of 3:1 (pink line panels A + C). To obtain the spectrum of the CH1 domain from the mix the individual CL or A33 VH domain spectrum was subtracted from the experimental mix (red/orange lines panels B and D), this was then compared to the spectrum of the CH1 domain in isolation which had previously been measured (yellow line panels B + D).

The spectrum of the CL domain cannot be observed as there is a change occurring to the spectrum of the mixed domains. Previous work carried out in the literature indicates this change is occurring in the CH1 domain. Furthermore as the CL domain is already folded and in an ordered state the change occurring is more indicative of the unfolded CH1 domain.

4.10 Discussion

The generation of engineered monoclonal antibodies for therapeutic use is a growing industry. Therefore, the production and characterisation of these therapeutic proteins is an important area of current research. The results presented in this chapter have described an alternative pathway for antibody folding and assembly which involves the secretion of HC dimers.

There have been cases of HC dimer secretion reported previously including camelids, hybrid immunoglobulins of sharks and various HC mutants discussed in the introduction. However, most of these HCs lack the CH1 domain which has been implied as the domain which controls antibody folding (Hendershot, Bole et al., 1987). The CH1 domain is essentially disordered being proline-rich and unable to form its hydrophobic core without the addition of LC. It forms a strong association with the ER resident chaperone BiP which binds to hydrophobic residues and retains the protein in the ER prior to assembly with the LC (Vanhove et al., 2001). We have demonstrated here that the CH1 domain can fold in the absence of LC. However, whether or not the CH1 domain exhibits its native structure is yet to be determined.

Other antibody domains have previously been reported to also have an important role in antibody assembly and secretion. As all immunoglobulin domains have a similar secondary structure, information regarding how the other domains fold may assist in understanding how the CH1 domain folds without the CL domain. For example, the CH3 domain of the HC which initiates dimerization of the Fc fragment also requires an isomerisation of a proline residue; however, the domain has been shown to fold prior to this reaction which is not the case for the CH1 domain (Thies et al., 1999). There is no interchain disulphide between the two CH3 domains, instead there is only an intrachain disulphide buried within the hydrophobic core of the each domain. However, it has been demonstrated that the two CH3 domains can fold and associate without this disulphide (Thies et al., 2002). However, in order to understand the folding and association of the CH1 domains, it is more appropriate to look at the folding of the CH2 domain as these dimerise with an interchain disulphide bridge between them as was seen with the dimerization of the truncated A33 Fab HC. The glycosylation site within the CH2 domain aids in stabilising the dimer, however un-glycosylated CH2 domains have been shown to associate and also require a proline isomerisation to fold (Feige et al., 2004). The folding and dimerization of the CH2 and CH3 domains of the HC are largely similar to that of the CH1 domain, however the stabilisation required seems to be internal rather than the addition of another domain as is seen with the CL acting on the CH1. It also seems to be the case that the CH2 and CH3 domains are folded prior to association with each other. Therefore, as the CH1 domain seems unable to fold in isolation, it could be the case that the opposite is true for the CH1 domain in that an interaction between CH1

and CL would allow the CH1 domain to fold. For the A33 HC the folding of the CH1 domain is variable region dependent, so two variable domains could be aiding in stabilising the CH1 domain and interacting with each other in order for the CH1 domains to interact and fold.

The results described in this chapter have implicated the variable domain in the folding of the CH1 domain; however this ability is not uniform across all VH domains. Of the seven HCs investigated, four were able to support secretion of HC dimers to varying degrees and three were not. Therefore, the reason for the ability of the VH domain to stabilise the CH1 domain must lie in the differences between the variable domains. As variable domain development and evolution is a complex process resulting in a numerous potential variable domains, characterising each one leads to difficulties when making comparisons or identifying trends amongst them. The stability of different variable domains varies greatly as does the sequences, mainly within the CDR regions, particularly CDR3. The stability of each variable domain contributes to its ability to fold and its susceptibility to degradation which are useful investigative parameters when designing and engineering variable domains for therapeutic use (Ewert et al., 2003). Investigations into the folding and assembly of variable domains has shown that the folding and association of the VH and VL domains can occur in only 30 seconds in the absence of the rest of the polypeptide chain, however, this process is concentration dependent as at higher concentrations the proteins tend to form incorrect structures (Hochman, Gavish et al., 1976). This quick association and folding of the VH and VL domains could potentially provide information as to how the VH domains can associate, therefore inhibiting association of the VL domain and causing the secretion of HC dimers in the presence of LC. The association of the VH domain with the LC has been shown to be inhibited by the introduction of a point mutation within the VH domain reducing the amount of IgG secreted, if this was used to instead inhibit association of the VH domain with itself it could potentially inhibit the secretion of HC dimers (Wiens, Lekkerkerker et al., 2001). Another example of the drastic effect of the introduction of a point mutation is the substitution of glutamine for glutamate at position 6 in the VH domain. This mutation resulted in lack of antigen binding and incorrect folding (de Haard, Kazemier et al., 1998). However, both glutamine and glutamate are conserved at position 6 in the framework of the VH domain, of the HCs investigated in this chapter HCs 4, 3 and 2 contain glutamine and HCs 1, 5, 497 and A33 contain glutamate. As both contain secretors and non-secretors of HC dimers this conservation does not seem to be attributing to this particular phenotype, however the study does highlight that subtle differences in the sequences of the variable domains can contribute to the overall structure and function, even potentially HC dimer secretion.

There are not any known cases of WT HC dimer secretion in mammalian cells, however, it has been reported that HC dimers can be secreted in the absence of LC in a *Drosophilla* cell line. This

was the first reported case of HC dimer secretion in invertebrates and in the study it was also shown that the HCs had the ability to bind to the BiP-like protein hsc-72 which contains 80% homology with human BiP (Kirkpatrick, Ganguly et al., 1995). This result correlates with data presented in this chapter that HCs can form dimers and be secreted via the association and disassociation of the chaperone BiP in the absence of LC. However this contradicts previous work where no HC dimer secretion was reported in the absence of LC in myeloma cells (Morrison & Scharff, 1975). Furthermore, interactions with BiP and secretion of HC dimers have been attributed to the cysteines present in the CH1 domain. When all cysteines within the CH1 domain of IgG were mutated, the HC was able to be secreted and did not associate with the chaperone BiP (Elkabetz, Argon et al., 2005). However, this is not applicable with the cysteine mutant of the truncated HC investigated in this chapter as regardless of the presence or absence of the disulphide bond present, HC dimer secretion still occurs. BiP association has not been confirmed with the truncated HC, but as it associates with the full-length HC dimer which is secreted, it can be suggested BiP is also associating with the truncated HC. This result highlights that the stable binding of BiP to the CH1 domain can be disrupted in the absence of LC.

In conclusion, this chapter has highlighted an additional role of the variable domain in its ability to contribute to the folding and secretion of HC dimers. This information can be used as a screening tool to identify potential IgGs which demonstrate the phenotype of HC dimer secretion. This could then allow for further engineering to avoid the phenotype so the immunoglobulin can be utilised for therapeutic use. For example, if the specific changes which are made from the WT protein were introduced more carefully, whilst screening for the phenotype of HC dimer secretion, any potential changes which are contributing to association of the VH domains can be avoided and altered. A site directed mutagenesis approach of antibodies which already display the phenotype could also be characterised to identify any trend observed with particular residues within CDRs or the framework. If the potential interactions of the VH domains were modelled this could provide a starting point to attempt to identify the residues responsible for the interaction prior to further analysis using techniques such as X-ray crystallography. If HC dimer secretion was hindered this might improve the yield of IgG produced and strengthen the association the VL domain has with the VH domain rather than the VH domain interacting with itself.

Chapter 5 : An investigation of the folding, assembly and secretion of FabFvs.

5.1 Introduction

The engineering and development of novel antibody formats is a relatively new and fast growing area of therapeutics with over 60 different formats in development (Spiess et al., 2015). The purpose of these therapeutics is to manipulate the strengths and weaknesses of immunoglobulins to engineer molecules with desirable effects. This manufacturing process has produced several undesired intermediates and by-products for different antibody formats. The greatest area of development and troubleshooting is the production of bi-specific antibody formats which are capable of recognising two separate antigen targets (Nisonoff & Rivers, 1961). The generation of bi-specific antibodies comprising of two separate HCs and LCs has proved difficult due to the production of homodimers. This issue has been overcome in several ways, one example is by mutating the CH3 domain of the antibody in a way to engineer “knobs and holes” in a lock and key type mechanism which allows dimerization of two different HCs to form heterodimers and inhibits homodimer formation (Ridgway et al., 1996). Another alternative way to engineer bi-specific antibody formats is to manipulate and mutate the Fv area of the antibody which is a single VL domain combined with a single VH domain. By investigating the VL/VH interface of a desired Fv, mutations have been made to either strengthen the H-bonding or electrostatic interactions to favour desired VL/VH conformations and pairing (Igawa et al., 2010, Tan et al., 1998).

One antibody format which has been developed by UCB Pharma is the FabFv (Figure 5.1). The structure of the FabFv consists of a single LC and a single HC which contain two variable domains and a single constant domain. The “Fab” comes from the VL and CL and VH and CH1 pairings seen in a complete immunoglobulin structure, and the “Fv” comes from a single VL and VH domain. The Fab and Fv structures are held together via a 15 aa long glycine and serine linker. The glycine-serine linker is very flexible which aids the therapeutic in binding to two separate antigen targets. Linkers are a common addition to antibody formats to enable additional domains to be inserted or to replace any domains which are deleted (Jakob, Edalji et al., 2013). FabFvs also contain engineered cysteine residues to allow an interchain disulphide bond to form between the second VL and VH domains to increase the stability of the molecule. This disulphide is in addition to the interchain disulphide which already exists between the CL and CH1 domains.

One advantage to the FabFv is that it lacks the Fc fragment of IgG. This can be a desired attribute as a therapeutic if the drug action designed requires a lack of binding to Fc receptors. The binding of IgG to their Fc receptors is a natural process in the immune response; however, it can cause antibody-dependent cellular cytotoxicity (ADCC) by activating the complement system. This can be useful for some drugs which target effector cells as seen with the bi-specific antibody anti-

EGFR x anti-HER3 in the treatment of cancer (Schaefer, Haber et al., 2011). However, if this activation is undesirable then the only way to completely inhibit it is to remove the Fc fragment. One disadvantage of the removal of the Fc fragment is that the half-life on the antibody format is short. Again, this could be a positive or negative attribute depending on the function of the therapeutic. One way to overcome the issue of a shortened half-life is to engineer the Fv fragment of the FabFv to bind to albumin. Albumin is a highly abundant protein present in the blood and would successfully extend the half-life of the FabFv (Tijink, Laeremans et al., 2008).

The development of the FabFv monomer has produced a successful antibody format. However, the assembly and secretion of these molecules produces some multimeric species of the FabFv (unpublished data by UCB Pharma). These multimers can cause a loss in the yield of monomer FabFv produced. The process of the folding and assembly of FabFvs is unknown. The results described in this chapter propose a pathway which drives multimer formation. The results suggest a potential mechanism to shift the equilibrium of assembly to the formation of monomer FabFvs.

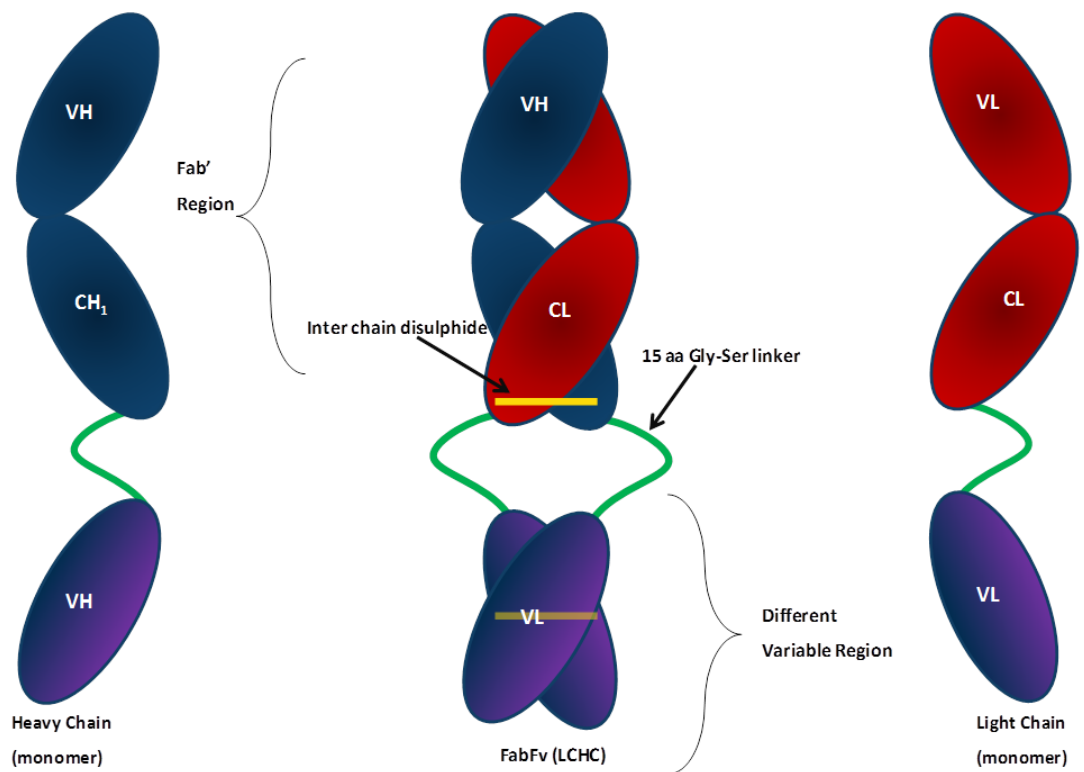


Figure 5.1: Schematic of the structure of FabFv.

The FabFv is a novel antibody format comprised of a single LC and single HC. Each chain is made up of three domains; one constant (CH₁ and CL) and two variable. The “Fab” fragment resembles that of full IgG with the variable and constant domains and the “Fv” fragment resembles the additional variable domains. The two regions are held together via a 15 aa long glycine and serine linker. The Fv region has an engineered interchain disulphide which adds stability to the proteins structure. The FabFv is a bi-specific antibody format developed by UCB Pharma.

5.2 Aim

The overall aim of this chapter is to characterise the assembly and secretion of FabFvs and understand how multimeric species of the molecule are forming. The approach taken to understand the assembly of the FabFvs involved carrying out single transfections of the LC and HC to observe the intermediates formed and secreted. This information was also combined with the studies of two chains expressed together to identify the different species which were assembled and secreted.

5.3 Identification of FabFv folding intermediates

In order to assess the intermediates and secreted products generated by FabFvs a similar approach was taken to the identification of HC dimers described in Chapter 4.2.

CHO cells were co-transfected with κ LC (497 FabFv) and IgG1 HC (497 FabFv) and incubated for 48h. A pulse chase assay was carried out over a time course of 5 hours with time points of 0, 1, 2, 3, 4 and 5 hours with an un-transfected negative control and a reduced sample both at 5h time points (Figure 5.2A + B). There are multiple protein bands which correspond to several FabFv species in the lysate and medium which are resolved upon addition of DTT. This implies the higher molecular weight species are held together via interchain disulphide bonds. There is also a protein band which remains unchanged overtime and is not affected by the addition of DTT migrating to approximately 75 kDa. The molecular weight of this protein suggests that it is the ER resident chaperone BiP which is known to associate with the unfolded CH1 domain (Bole et al., 1986) and still associates with this antibody format.

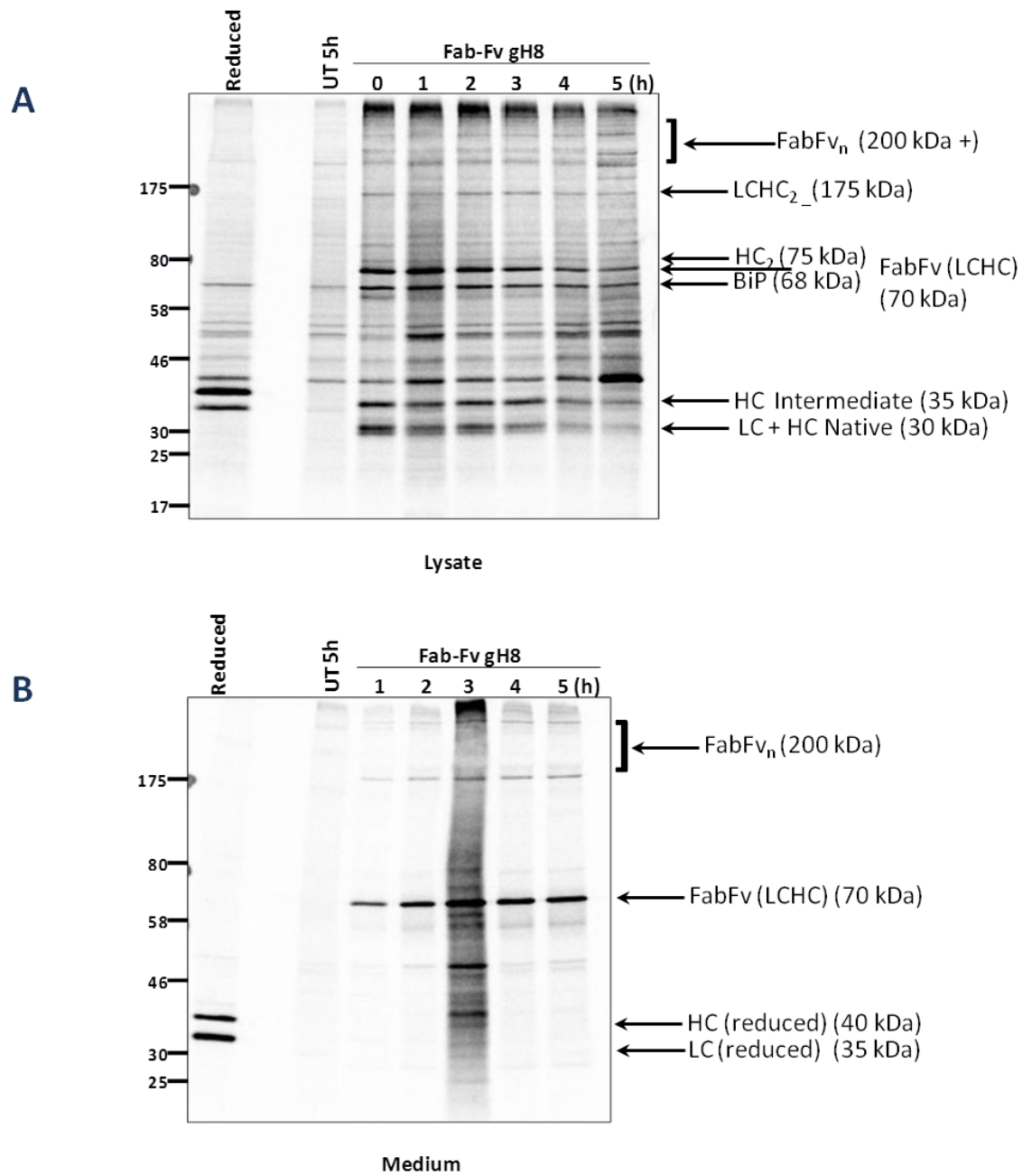


Figure 5.2: Assembly and secretion of FabFv 497 over time.

CHO cells were co-transfected with equal amounts of LC and HC. After a 48h incubation period cells were pulsed with ³⁵S methionine/cysteine labelled medium and chased with label-free medium for up to 5 hours with time points of 0, 1, 2, 3, 4 and 5 hours. FabFv products and intermediates were isolated and run under non-reducing conditions with a single reduced time point plus an un-transfected negative control on a gradient (4-20%) polyacrylamide gel. The products within the lysate (A) and medium (B) have been identified.

In order to identify the composition of the multimer species being formed, 2D gel electrophoresis (Chapter 2, section 2.3.3) was used with a first dimension non-reducing and a second dimension reducing (Figure 5.3).

From the initial pulse chase assay time course, it was deduced that an optimum chase time of 3 hours was sufficient to visualise the contents of the protein bands in both the lysate and the medium (Figure 5.2A +B). CHO cells were transfected and radiolabelled as described previously, the lysate and media were collected and the FabFv products were isolated using protein A Sepharose beads plus anti κ LC antibody to isolate any LC products and/or intermediates. The Fc fragment is usually required for binding to protein A; however the FabFvs engineered by UCB Pharma have the added attribute of protein A binding for isolation and purification.

There are several intermediates and complete FabFv which can be identified following 2D analysis (Figure 5.3). Firstly, there are two protein bands in the cell lysate which migrate to approximately 30 and 35 kDa in the non-reducing dimension (highlighted in red (square) Figure 5.3). Analysis of these two protein bands in the reducing dimension revealed that both represent the HC of FabFv. This indicates that the HC exists in two forms; the first upper band in the non-reducing dimension is a folding intermediate of the HC, most likely with its CH1 domain disulphides not formed and the second form of HC which migrates faster through the gel in the non-reducing dimension is the HC containing all its correct disulphides. Secondly there is another product within the lysate which only contains HC that migrates to approximately 75 kDa in the non-reducing dimension (highlighted in pink (star) Figure 5.3). Due to its molecular weight in the non-reducing dimension and the fact that it only contains HC in the reducing dimension, it can be deduced that this intermediate is a HC dimer.

The other protein bands within the cell lysate contain both LC and HC; however, there are two distinct bands in the non-reducing dimension migrating to approximately 100 and 70 kDa (100 kDa = Y and 70 kDa = X in the non-reducing dimension) that seem to differ in the intensity of LC and HC in the reducing dimension (highlighted in green (circle) Figure 5.3). The protein bands representing HC and LC in the reducing dimension were quantified. The ratio of intensity of LC to HC within the species X and Y in the non-reducing dimension was then determined. It was found that band X (70 kDa non-reducing) comprised of equal amounts of LC and HC in the reducing dimension indicating a single LC and HC therefore representing the monomer of FabFv. However, the Y protein band (100 kDa non-reducing) was found to contain twice as much HC as LC in the reducing dimension, therefore identifying this species as a folding intermediate comprising of a HC dimer with a single LC (LHC₂).

In contrast to the different intermediate species within the lysate, the components of the protein species found in the medium all consist of both LC and HC in equal ratios (Figure 5.3). From the different molecular weights presented in the non-reducing dimension it can be seen that as they all contain equal amounts of LC and HC these secreted proteins represent the monomer and multimer forms of FabFv. However, in the medium there is a species in the non-reducing dimension that migrates to approximately 30 kDa which corresponds to the size of a single HC when compared with the protein in the lysate (compare red square and purple triangle Figure 5.3). However, the 30 kDa protein in the medium contains both HC and LC in the second reducing dimension (highlighted in purple (triangle) Figure 5.3). Due to its molecular weight and the fact that both LC and HC are present, one possibility for this observation is that this species is a FabFv which has not formed its interchain disulphide. If this is true, this 30 kDa protein represents a secreted FabFv held together via non-covalent interactions. However, the single monomer chains could also be associating with either multimeric species of FabFv, themselves as homodimers (LC₂ or HC₂) or with nothing at all (secreted as single chains). In conclusion, it can be said that the single protein band seen in the non-reducing dimension actually represents both the LC and HC of the FabFv, however as the conditions of the gel are denaturing the non-covalent single chains in their native state are still unknown.

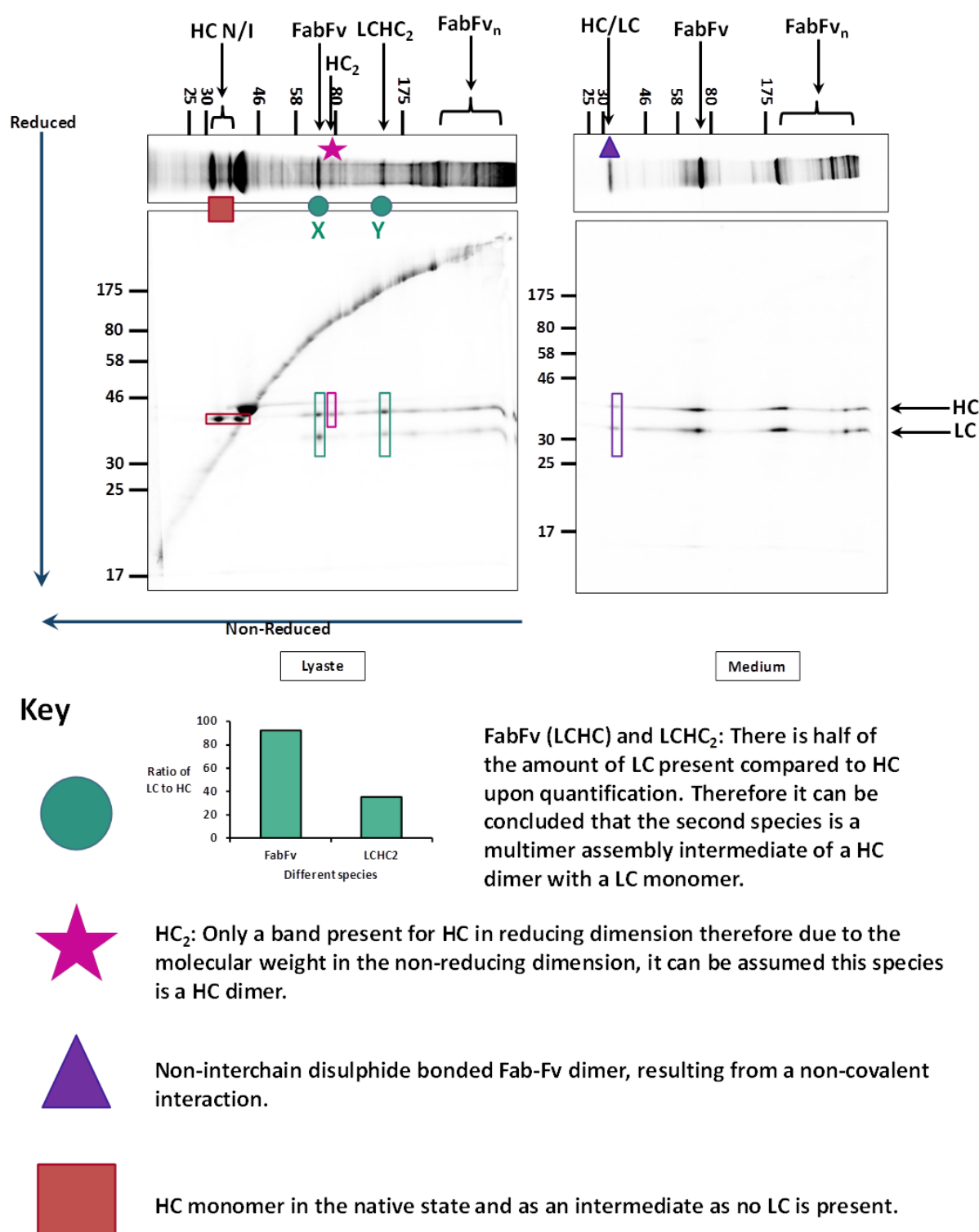


Figure 5.3: 2D gel analysis of FabFv folding intermediates and secreted products.

CHO cells were co-transfected with equal amounts of LC and HC (FabFv 497). Cell lysates and media samples were collected and prepared via immunoprecipitation with protein A Sepharose (plus anti κ LC). Cells underwent a 3h pulse chase assay and a 2D radiograph was produced of intracellular (left) and extracellular (right) material of FabFv 497. The quantification of the radiolabelled protein bands corresponding to LC and HC are comparable as each chain contains an equal amount of methionine and therefore will have incorporated equal amounts of radiation.

5.4 LC FabFv expressed alone results in multimer formation and secretion

After analysis of the intermediates formed and products secreted from co-transfecting LC and HC FabFvs, the expression of the single chains was investigated to characterise further the different multimeric species produced.

CHO cells were transfected with κ LC FabFv 497 and incubated for 48h. A pulse chase assay using ^{35}S cysteine/methionine radiolabelled media was undertaken for a pulse time of 30 minutes and a chase time of 60 minutes with time points at 0, 5, 30 and 60 minutes. The lysate and medium were separated and any LC products were immunisolated using anti κ LC incubated with protein A Sepharose beads. The protein samples were denatured and run on a gradient (4-20%) polyacrylamide gel with a reduced sample at 60 minutes and an un-transfected negative control (Figure 5.4A +B). There are several higher molecular weight species within the lysate and medium of the LC transfected cells which are resolved upon addition of DTT indicating they contain interchain disulphide bonds. Within the lysate of the cell, two protein bands which are migrating to approximately 35 kDa and are not present in the un-transfected negative control can be deduced as a partially folded LC monomer and the fully folded monomer. There are a further two species which can be identified in the non-reducing lanes over the time course which are specific to LC as they are not present in the un-transfected control. The molecular weight of each protein band corresponds to approximately 60 and 80 kDa. From their molecular weight these two complexes could be a LC dimer and a LC trimer which are held together via interchain disulphide bonds. Upon comparison with the secreted material, it can be seen that all three species are secreted after 1 hour of chase time (Figure 5.4B). Our proposed model for LC assembly and secretion indicates that a partially folded LC forms its native state and forms a dimer and/or trimer before being secreted (Figure 5.4C). The monomer form of the LC which is seen to be secreted however may exist as a dimer or trimer in solution via non-covalent interactions as potentially seen with the LC and HC of FabFv in the 2D gel analysis of the secreted material (highlighted in purple (triangle) Figure 5.3B).

Therefore it can be seen that the folding and assembly of the LC in the absence of HC shows similarities with the two chains expressed together as in both cases multimeric species are formed and secreted.

5.5 HC expressed alone results in dimer formation and secretion

Investigations of HC alone transfections previously described throughout the results chapters have demonstrated that HC dimers with specific VH domains can fold and be secreted in the absence of LC. It was also described in the previous chapter that the Fc fragment of the IgG protein is not required for dimerization or secretion. As FabFvs are more complex with an additional Fv domain, two HCs with the N terminal VH domain known to secrete HC dimers were investigated.

Initially, the HC counterpart of FabFv 497 was transfected alone into CHO cells and incubated for 48h. A 5h pulse chase assay was carried out with a reduced sample and an un-transfected negative control, with time points of 0, 1, 2, 3, 4 and 5 h (Figure 5.5A +B). The lysate and medium were separated and the proteins were isolated using protein A Sepharose beads. Within the cell lysate in the non-reducing lanes over the time course there are three protein bands specific to HC; HC dimer, HC monomer partially and fully folded. These intermediates can be easily identified as they were seen with the co-transfection with LC in the 2D gel analysis (Figure 5.3). The protein band resembling HC dimer is also resolved in the reducing lane with only a single band representing HC. BiP is also present with samples that contain HC FabFv 497 which is expected as BiP has a known association with the HC. However, secretion of HC dimers occurs later than seen with the LC at a chase time of 2h.

To compare this result with another FabFv HC with a different N-terminal VH domain, the HC A33 FabFv was transfected into CHO cells and incubated for 48h. A 5h pulse chase assay was carried out with a reduced sample and an un-transfected negative control, with time points of 0, 1, 2, 3, 4 and 5 h (Figure 5.5 C +D). The results of the pulse chase assay show that the same intermediates can be identified in the lysate and medium samples as seen with HC 497 FabFv, again with secretion of HC dimers in the medium samples occurring after a time point of 2 h (Figure 5.5D). However, the time point seen for the secretion of HC dimers differs from the full length HC which is secreted after 1h for both A33 and 497 VH domains. In addition, there is evidence of multiple disulphide species of HC dimer which are secreted and are not seen for the full length or Fab HC counterparts. Therefore, the proposed model for HC assembly and secretion for FabFvs is that a partially folded HC with an unfolded CH1 domain folds via the same mechanism proposed in Chapter 4, forms a dimer and is secreted after a time point of 2 hours (Figure 5.5E).

An interesting observation of these HC only pulse chases is that the only multimer formation occurring is dimerization which is seen with all HCs of IgG. There are no other higher molecular weight multimer formats observed with the HC only transfection. This could imply that the

interactions between the two chains of HC dimer are more stable than the LC interactions and therefore multimer formation does not occur.

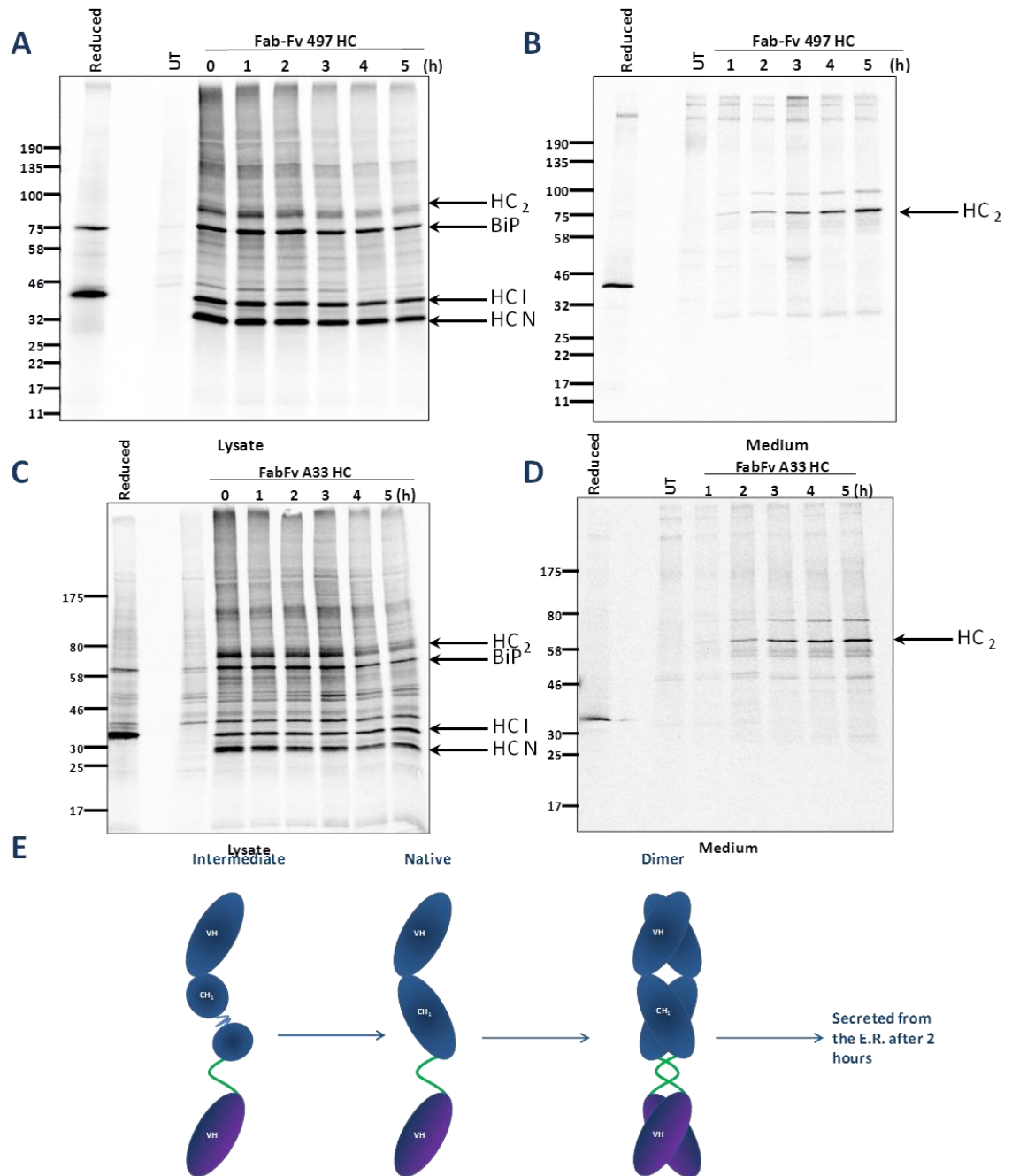


Figure 5.5: Assembly and secretion of FabFv HC.

CHO cells were transfected with the HC of 497 (A + B) and also separately transfected with the HC of A33 FabFv (C + D) and underwent a pulse chase assay with time points of 0, 1, 2, 3, 4 and 5 hours. The lysate and medium show the formation and secretion of HC dimers after 2 hours. The proposed model for HC FabFv assembly and secretion can be seen in panel E.

5.6 Summary of initial intermediates formed and secreted during FabFv assembly

In order to gain a clear picture of the intermediates and products formed and secreted during the initial synthesis of FabFvs the various species produced were compared on the same gel.

CHO cells were separately transfected with either κ LC FabFv 497, HC FabFv 497 or both chains co-transfected together and incubated for 48 h. Each transfection underwent the pulse chase assay previously described for a single chase time of 60 minutes. The samples were then denatured and underwent gel electrophoresis in both reducing and non-reducing conditions (Figure 5.6). Each medium sample was run next to its corresponding lysate sample to observe which bands were secreted. The results in the non-reducing conditions mirror that of what has previously been described. It is clearer from this result that the fully disulphide bonded LC and HC under non-reducing conditions migrate to the same place on the gel (compare lanes 1 and 3 Figure 5.6A). This result gives an explanation as to why only one protein band can be visualised in the medium sample on the 2D gel analysis yet in the reducing dimension both the LC and HC are present (highlighted in purple (triangle) Figure 5.3B). The HC and LC migrate to separate molecular weights under reducing conditions (compare lanes 1 and 3 Figure 5.6B), the differences seen with the migration of the chains in the presence or absence of DTT can be attributed to the intrachain disulphides.

After 1 hour of chase time there still seems to be no secretion of HC dimers in the absence or presence of LC, whereas the LC expressed alone still secretes in multimeric forms, as does the FabFv. These results demonstrate that the intermediates formed within the FabFv assembly pathway and the ability for the FabFv 497 HC to be secreted is dependent on the presence of LC.

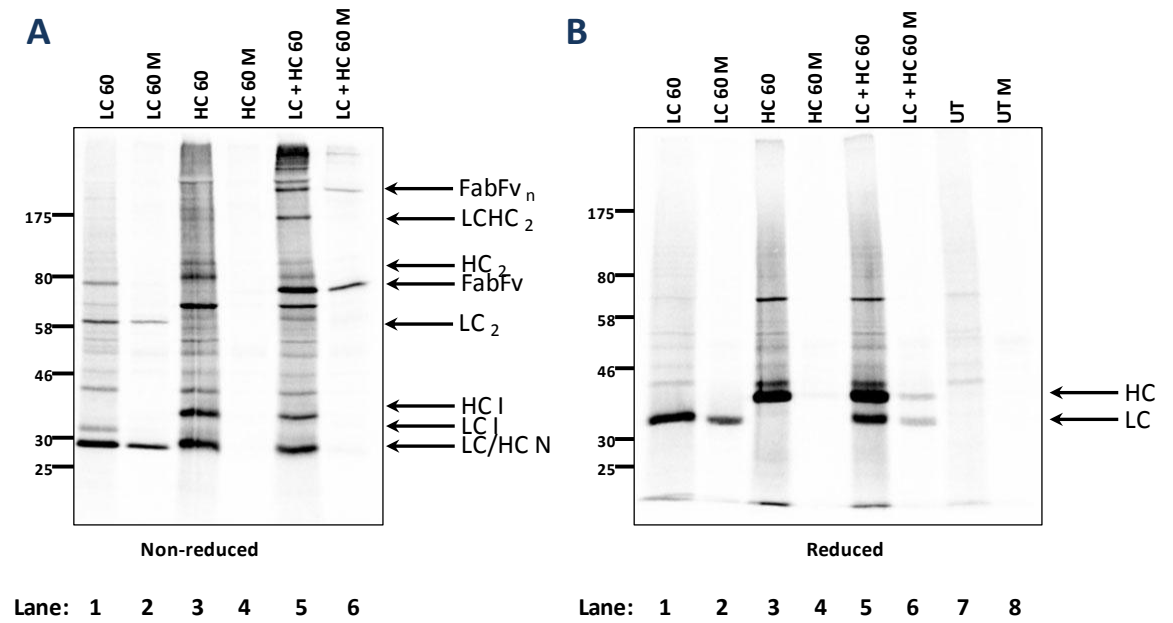


Figure 5.6: Summary of FabFv assembly and secretion.

CHO cells were transfected with either κ LC FabFv 497, HC FabFv 497 or both chains co-transfected together and pulse chased for a time point of 60 minutes. The lysate and medium samples were run next to each other either under non-reducing (A) or reducing (B) conditions. The intermediates formed and products secreted in each condition are labelled giving a clearer picture of the differences between the chains assemblies.

5.7 Discussion

The engineering and production of novel antibody formats is a growing area of research in the production of biological molecules which can be used as therapeutics (Chan & Carter, 2010). The manipulation and re-design of antibodies allows for the design of highly specific drugs which include the actions of the human's natural immune response. However, as these antibody formats are complex and novel proteins, the interactions which occur during their folding and the unwanted products which can assemble are unpredictable, requiring further manipulation to increase the yield of the desired product designed for therapeutic use.

FabFvs are a unique antibody format developed by UCB Pharma where they have combined the effect of a bi-specific antibody lacking the Fc fragment but increased the stability. Single-chain Fv (scFv) antibody formats are widely reported as a therapeutic as this is the portion of the antibody which is involved in antigen binding. The small protein is useful as it increases the ability of tissue penetration that cannot be carried out by a full length IgG. However the scFv is by itself less stable than the Fab fragment in terms of rate of denaturation, as it does not yield the same chain interactions or disulphide bond arrangement which brings the CL and CH1 domains together to give the Fab fragment its increased stability (Rothlisberger et al., 2005). The FabFv has overcome this instability problem through engineering a disulphide in the Fv part of the protein while linking it to the more stable Fab fragment. However, from the results presented in this chapter, the second engineered interchain disulphide is implicated in the formation of multimeric species of the FabFv.

The results presented in this chapter have provided information to propose a model for the formation of multimer species during FabFv assembly (Figure 5.7). To summarise the evidence gathered thus far, we show that several intermediates are involved in the FabFv assembly pathway; LC and HC folding intermediates, HC dimer, FabFv monomer (LCHC), LCHC₂ and multimers (FabFv_n). The intermediates formed after 1 hour for each chain alone are LC monomer, dimer and trimer (all secreted) and HC monomer and dimer. However, there is secretion of HC dimers seen after 2h of chase time. There are two protein bands present in the medium after 2h which represent HC as that is the only species present in the reducing lane of the gel. Although the HC dimer present in the medium migrates to the same size as BiP in the lysate, the lack of BiP in the reducing lane confirms that it is the HC which is being secreted. The multiple forms seen in the medium could be alternative species of HC dimers. As the HC can form two potential interchain disulphide bonds; one within the CH1 domain and one in the C-terminal VH domain, if only one has formed during synthesis this could explain the two HC dimer bands seen upon secretion of the HC. In order to determine if this is the case, one could mutate the individual

cysteines which are involved in the interchain disulphide bonds to alanine to determine and observe which species are secreted to confirm the disulphide status of each HC dimer. Furthermore, our results suggest that the HC dimer which has formed its intrachain disulphide within the CH1 domain is migrating at the same place on the gel as BiP and therefore cannot be clearly identified in the lysate. However, HC dimers are not seen to be secreted at all when co-expressed with the LC which could be due to the slower folding and secretion seen with the HC expressed alone.

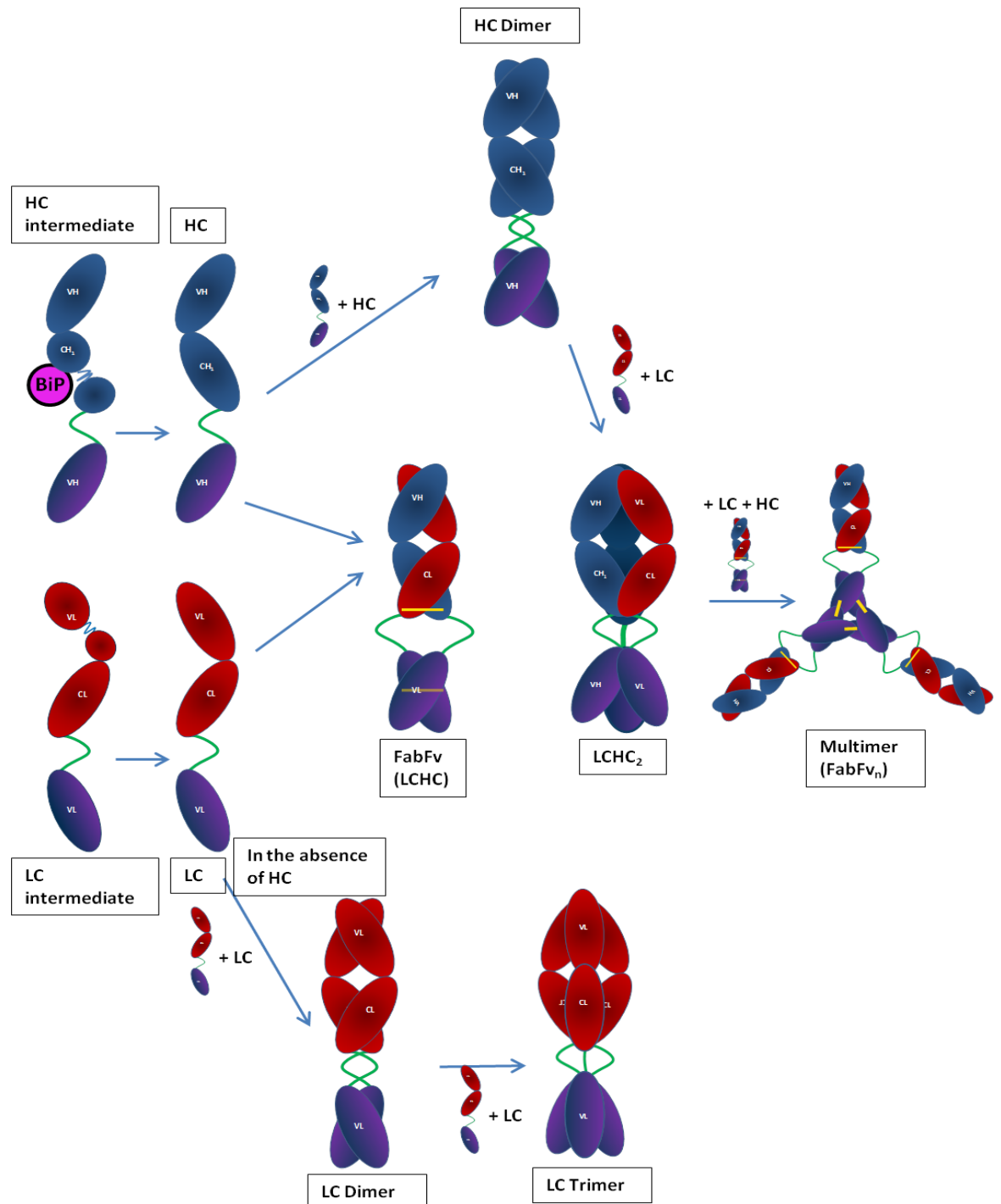


Figure 5.7: Model of multimer assembly of FabFvs.

The proposed model for FabFv folding and assembly follows that the LC and HC are both translocated into the ER where they begin to fold noted by the intermediates, and then form folded single chains. HC expressed alone form HC dimers which are secreted after 2 hours; this intermediate is also seen in the presence of LC but is not secreted. When the LC is expressed alone it forms multimer species which are all secreted only in the absence of HC. When the two chains are co-expressed together, several intermediates form including the single individual chains, a HC dimer, a FabFv (LCHC), a LCHC₂ and multimers. It has been hypothesised that the multimer species of the FabFv is orchestrated via the LC. The LCHC₂ could be an assembly intermediate in the multimer formation pathway. Alternatively, alongside this assembly the FabFv monomers may also be assembling into multimers via the second VL domain of the LC.

As multimer formation is seen following co-transfection of LC and HC and when LC is transfected alone, this indicates that the association of the chains may be via the LC (Figure 5.7). There are several similarities between the assembly and secretion of the LC and the FabFvs; firstly both form multimeric species which are secreted, secondly both secrete what appears to be a single chain. In the case of the LC expressed alone, it can be seen that the LC monomer is secreted, however when LC is co-expressed with HC both the LC and HC monomers are secreted migrating to the same place on the gel under non-reducing conditions. It can be suggested that the apparent monomer secretion of LC and HC is actually the secretion of a FabFv which has not formed its interchain disulphide bonds and is interacting non-covalently prior to denaturation for gel electrophoresis. If this interaction is correct, it can also be hypothesised that the LC monomer which is secreted when LC is expressed alone exists as a dimer or trimer which has not formed its interchain disulphides but instead is interacting non-covalently. When LC and HC are co-expressed together there are no singular (LC only or HC only) species which are secreted detected in the 2D gel electrophoresis. This evidence in conjunction with the hypothesis that the FabFv chains are able to associate through non-covalent interactions implies that LC has a preference to associate with HC rather than itself as no multimeric species of solely LC are present when the two chains are expressed together.

Other evidence that the LC may be governing multimer formation is the presence of a LCHC₂ species identified following 2D gel analysis. HC dimers have been seen to form when HC is transfected alone and when co-transfected with LC but they are secreted at a slower rate, and they do not form any higher multiple species other than the dimer. However, identification of the LCHC₂ species suggests that once the HC dimer is formed, LC can associate with HC dimers as no multimers of the LC can be seen following co-transfection with HC. The association of LC with the HC dimer again suggests that LC has a preference to associate with HC rather than itself. Other LCs/FabFv monomers could continue to associate via the LC which would result in the multimer species which are formed. However, if multimer formation is occurring through the LC, it is not clear which domain is controlling it as there are two available cysteines, one within the CL domain and one in the second VL domain, both of which are capable of potentially forming disulphides. It is unclear from the information gathered thus far whether multimer formation is occurring through domain association or disulphide formation. However, as the multimeric species do contain interchain disulphides, if domain association is the driving force behind multimer formation, it can be said that the interactions are stabilised through disulphide bonds. In order to determine through which domain multimer formation is occurring, one must investigate each potential disulphide separately. If the Fv fragment was expressed alone with or without the potential to form a disulphide then this could answer whether or not this part of the FabFv is

responsible for multimer formation. Furthermore, mutating the cysteine within the second VL domain to alanine to inhibit the potential for an interchain disulphide could also potentially implicate its role in multimer formation. However, as the interchain disulphide bond between the C-terminal variable domains is due to engineered cysteines, and multimer formation is not seen with the LC of IgG which lacks the additional VL domain (Chapter 4.5) this strengthens the possibility that the multimer orchestration is via this additional domain.

Chapter 6 : Discussion

6.1 Overview

The results presented in this thesis have addressed the question of how to improve the yield of therapeutic antibodies and novel antibody formats. Several anomalies in the assembly and secretion of these antibodies have been identified. Firstly, the cleavage of the HC A33 IgG4 was demonstrated. This cleavage event has highlighted the potential to screen for cleavage of HCs for therapeutic use. If this process could be inhibited, it would potentially increase the yield of the antibody if the HC is no longer being cleaved. Secondly, the light-chain independent secretion of HC dimers has been shown to occur for some engineered antibodies which, if avoided would increase the yield and homogeneity of IgG produced. Thirdly, our results propose a cause of multimer formation of FabFvs which suggests a mechanism to inhibit this process with a view to increase the yield of the therapeutic monomer produced. This chapter will discuss the results presented outlining any shortcomings, in addition to putting forward new ideas for further improvements.

6.2 How and why is the HC A33 IgG4 being cleaved?

Chapter 3 describes the investigation of the A33 IgG4 antibody the HC of which is cleaved resulting in a single Fc fragment produced. The mass spectrometry data analysis of an unknown 30kDa protein band yielded three peptides which corresponded to the CH2 and CH3 domains of IgG4. As the cleavage product appears as a non-disulphide-bonded monomer of the Fc fragment, this indicates that cleavage is occurring C-terminal to the hinge region of the protein; however the protein could potentially dimerise via non-covalent interactions. One peculiar observation, however, is that this cleaved part of the protein remains unchanged overtime (longest investigation of 5 h) and the rest of the cleaved protein does not seem to be present, however, this is most likely attributed to the inability of the Fab fragment to bind to protein A.

Comparison of A33 IgG4 and A33 IgG1 showed that of the two isotypes the cleavage product was only present for IgG4. With this observation, two mutants of IgG4 were made; one which was the WT of IgG4 as all IgG4 HCs used in this thesis have a point mutation in the hinge region to avoid half molecule formation and an IgG4 which had the hinge sequence of IgG1. However, both mutants still proceeded to exhibit the cleavage product. Therefore, the next line of investigation in comparing the two isotypes should be to swap the CH1 domains as this domain is not present in the cleavage product, and observe if any cleavage still occurs. Within the amino acid sequence of the CH2 domain between the hinge region and the first peptide result from the mass spectrometry data there is one amino acid difference between IgG1 (leucine) and IgG4 (phenylalanine). As the cleavage product is only observed with IgG4 (with the variable domain A33) this amino acid difference could also be a potential site to mutate to determine if this is the

site where cleavage is occurring as swapping the hinge regions of IgG1 and IgG4 did not result in lack of cleavage of the HC.

Six other IgG4 HCs were then compared to A33 IgG4 HC to determine whether this effect was common to all IgG4 isotypes but no other HC demonstrated this phenotype. This would suggest that the cause of the HC cleavage is variable domain dependent as that is the only difference between the seven IgG4 HCs. However, as it is the same sequence of VH domain for both A33 IgG4 and A33 IgG1, then this cannot be the sole cause of the cleavage of the HC. In this particular case it seems to be a combination of variable domain and isotype which sets A33 IgG4 apart from the other HCs and perhaps this specific combination is causing cleavage. However, only seven HCs were tested which is not a good representation of the IgG4 isotype as the potential quantity of different variable domains range is approximately 10^6 - 10^7 . That being said, although this representation is small, the result suggests this anomaly is specific to A33 IgG4 HC. To be completely confident however, a larger number of HCs should be investigated. Also, it only seems to be a portion of the total HC protein in the ER to which this is occurring as the rest of the HC exists as an intact full length monomer, dimer or complete immunoglobulin. In addition, the cleavage product and full length HC both co-exist in the ER but only the full-length HC is glycosylated even though both proteins contain the same glycosylation site. One explanation for this selective glycosylation would be if the glycosylation is being blocked in some manner during the cleavage of the protein or a lack of glycosylation itself is causing cleavage of the HC which is inhibited upon glycosylation. One approach to test the hypothesis that a lack of glycosylation is causing cleavage would be to mutate the glycosylation site within the CH2 domain of the HC and observe whether inhibiting glycosylation increases the level of cleaved product. As the protease responsible for cleaving the HC is unknown, one could treat the transfected cells with a range of cell-permeable protease inhibitors during pulse chase analysis to try and define the type of proteases responsible by looking for the absence or presence of the cleavage product. Furthermore, as the amount of cleaved protein seems to remain unchanged, this creates two potential scenarios; either the protein is being cleaved and degraded at a steady state or not degraded at all. To test this one could use an inhibitor for ERAD such as MG132 during transfection and pulse chase analysis to observe whether the amount of cleaved HC increases which would suggest a steady turnover of the cleavage product instead of it not being the subject of degradation at all.

A useful tool to ultimately determine where within the sequence of the HC cleavage is occurring would be to carry out a more detailed mass spectrometry approach or to carry out N-terminal sequencing of the peptide. Why this particular HC seems to be susceptible to cleavage above other IgG4 HCs and its IgG1 isotype still remains unclear. However, the identification of the

cleavage product has provided an opportunity to screen potential therapeutic HCs to avoid a cleavage event and therefore increase the overall yield of the antibody produced.

6.3 Applications and causes of HC dimer secretion

Chapter 4 of this thesis describes the process of light chain independent secretion of HC dimers. Previous investigations into antibody assembly and secretion state that the HC remains in an unfolded state with a disordered CH1 domain sequestered in the ER via the chaperone BiP until the LC displaces BiP and is able to fold the CH1 domain to form a complete folded IgG (Feige, Hendershot et al., 2010). There have been reported cases where the LC and HCs eluded this quality control and were secreted without their partner chain (Mains & Sibley, 1983, Minegishi & Conley, 2001). For HC, these exceptions have occurred due to a lack of the CH1 domain or mutations within the HC (Hamers-Casterman et al., 1993, Muyldermans et al., 1994). The results reported in this thesis have described a mechanism for HC dimer secretion of humanised IgG which has not been seen before.

The secretion of HC dimers was suggested by the presence of two protein species in the medium from cells transfected with HC and LC; IgG and what was suspected to be a HC dimer. However, one possibility for this observation could be that the HC is folding with the aid of the LC with the lack of the interchain disulphide between the CL and CH1 domains. If this was the case the two chains could dissociate following denaturation and when separated on a gel appear to exist as a HC dimer with no LC. This hypothesis was incorrect however as demonstrated by the control experiments carried out in section 4.4, which included evidence that the folding of the CH1 domain was a prerequisite for HC dimer secretion, that the HC dimer had been processed in the GA prior to secretion and that the protein species secreted that was suspected to be a HC dimer was comprised of only HC. As all the experiments used to investigate antibody secretion were in a transient system, a good comparison would be to look for the presence of HC dimers in a stable cell line, potentially overexpressing LC to observe whether this has any ability to reduce the level of HC dimers secreted. For example, one investigation showed increasing the levels of LC in a stable system decreased the amount of aggregates secreted and formed intracellularly (Ho et al., 2015).

The HCs investigated in this PhD are engineered from a non-human source. As antibodies cannot be ethically produced in humans as a therapeutic, immunisation of animal alternatives are used to engineer a humanised form of an antibody for use as a drug. This process essentially produces a manufactured HC which is not naturally occurring. This observation provokes the idea that if HC dimer secretion is a naturally occurring process, there could be an unknown mechanism during B cell development which selects for HCs which are unable to secrete without a partner LC. For

example, it has been shown that HC dimer secretion is a variable region dependent process which can be reduced via mutations in the VH domain. Therefore, during B cell development at the pro B cell stage, where it has previously been reported that IgM HCs have been displayed on the cell surface, there could be a mechanism which would inhibit these cells from processing through B cell development to the pre B cell stage. This inhibition would avoid the generation of antibodies whose HCs undergo LC-independent secretion (Minegishi & Conley, 2001). In order to address this hypothesis, a wider repertoire of HCs must be investigated to identify a larger group of HCs which have the ability to secrete dimers. A Jurkat T cell line which does not express any endogenous LC or HCs could be used to screen for potential HC dimer secretors by transfecting with a library of potential HCs and then identifying any which are expressed on the cell surface using antibodies which recognise the Fc fragment of the HC for detection on intact cells (Costa, Franke et al., 1992). The variable domains of these HCs can then be sequenced from the cells expressing HC on the surface providing a much larger pool of variable domains capable of LC independent HC dimer secretion. With this information, potential sequence similarities or VDJ identities will be more reliably identified compared to the small pool of seven HCs investigated in this thesis. Once a sequence has been identified, this could be used as a screen to identify potential HC dimer secretors for use as a therapeutic prior to production. It could also provide information about the negative feedback mechanism which naturally inhibits the development of HCs which undergo LC independent secretion. Alternatively, as an initial test, to narrow down the area of the VH domain which is causing stabilisation of the CH1 domain the CDR3 regions of a secretor and a non-secretor could be swapped. CDR3 is the most variable part of the variable domain as it has the most interactions with the antigen. The amino acid length of this region is also highly variable as seen with the HCs investigated in this thesis; the secretors have shorter CDR3's than the non-secretors. This difference in length of CDR3 could be contributing to the association of the VH domains as spatially, it would provide a less protruding structure therefore potentially facilitating interactions between the two domains. If a longer CDR3 region of a non-secretor was swapped with a shorter CDR3 region of a secretor and a reduction or lack of HC dimer secretion was observed, this would implicate this area of the VH domain in stabilising the CH1 domain and therefore aid in narrowing the focus of sequence similarities to the CDR3s of new potential therapeutics. In addition to this investigation, the results of this thesis also showed a difference in levels of secretion of HC dimers among the secretors. Therefore, within this small pool, CDR swaps or a site-directed mutagenesis approach may also provide more information for any specific residues which are contributing to the association of the VH domains. For example, within the sequence of CDR2 at amino acid position 65 there is a conserved lysine residue for VH domains A33, 497 and HC3 but is glutamic acid for HC4. Small differences like this one can be

investigated to observe whether they are contributing to the secretion of HC dimers. If the CDR regions are implicated however, there is the potential that the resulting antibody would lose the ability to bind to its target antigen. If the CDR3 region is tampered with to a large extent this could result in a loss of binding function and would be no use as a therapeutic. Therefore, any mutations made would need to control for a balance of binding capacity combined with a lack of HC dimer secretion phenotype.

If it is not possible to identify sequence similarities which influence the secretion of HC dimers, then the underlying mechanism within the protein may provide an alternative line of investigation. From the results presented in this thesis, it has been demonstrated that an isolated CL domain can introduce structure into an isolated CH1 domain when the two are combined. This assay has been demonstrated previously, the results showed that the structure which was introduced into the CH1 domain was more pronounced than the findings in this thesis (Feige et al., 2009). The reason for this could be due to a number of possibilities. For example, the mix of the CL and CH1 domains in the study carried out by Feige and colleagues used an oxidised CL domain with its disulphide formed, whereas the purified CL domain described in this thesis did not undergo additional modifications before incubating with the CH1 domain. This lack of an oxidising environment could contribute to the rate of folding of the CH1 domain which could explain why after the same incubation time the structure introduced was less pronounced. Another possibility could be that the CL domain in the purified sample could be a mix of unfolded and folded protein as the spectrum from the CL data would only report on the folded structure. Therefore, if a small amount is not fully folded this would mean that the ratio of folded protein would not be as accurate as initially thought and may take longer to fold the CH1 domain. To determine if either of these are the case, higher ratios of purified protein could be used under oxidising conditions to provide the correct conditions for the CH1 domain to fold in the time frame presented by Feige and colleagues. This would also need to be repeated with the A33 VH domain as a control. However, the results demonstrated in this thesis show that the VH domain is unable to introduce structure into the CH1 domain but the CL domain does. Therefore, the mechanism behind the CH1 folding in the absence of the LC must be via an alternative action than is seen with the CL domain. The current working hypothesis of LC independent HC dimer secretion is outlined in Figure 6.1.

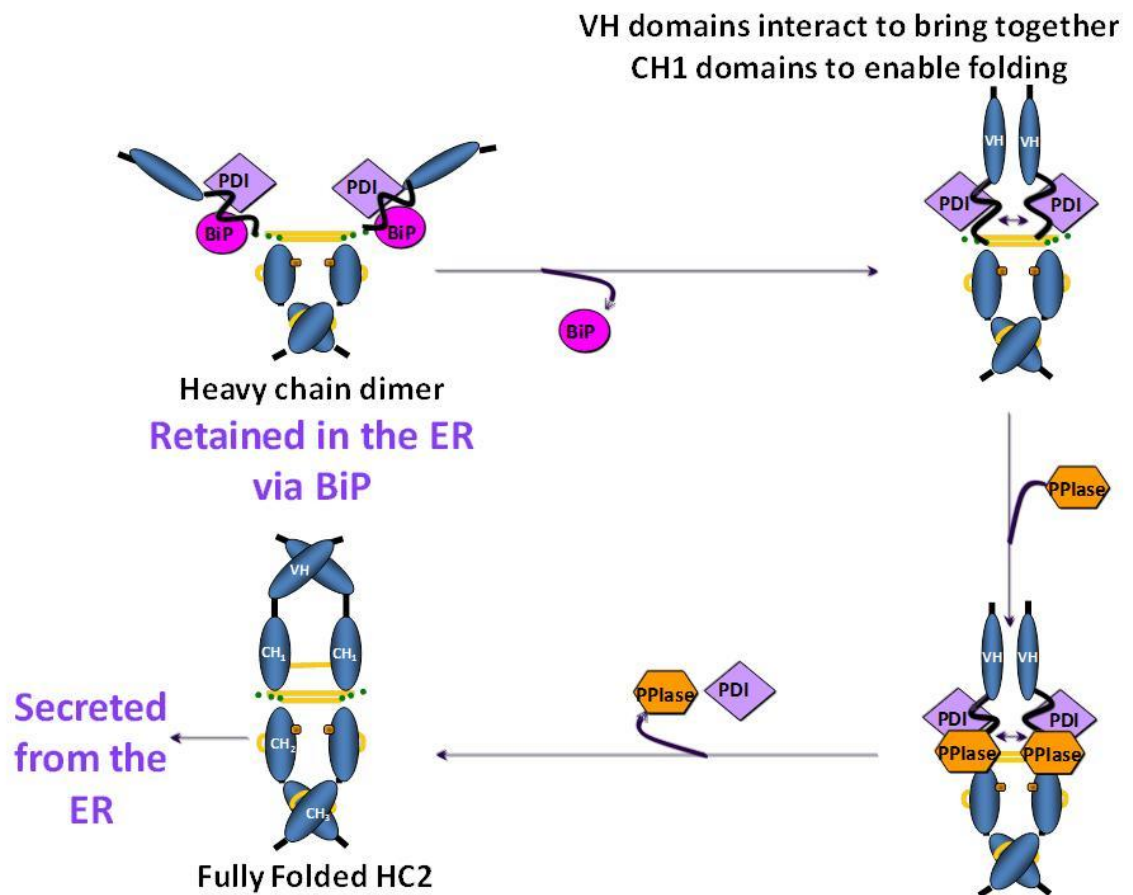


Figure 6.1: Hypothesis for the folding, assembly and secretion of HC dimers.

The HCs (blue) of IgG co-translationally enter the ER with each domain beginning to fold. The CH1 domain is the last to fold and remains in a disordered state and retained in the ER via the chaperone BiP (pink). Normally, LC would then displace BiP; however in the absence of LC, it has been proposed that some VH domains can interact with each other in the same manner as the VL and VH domains. This interaction enables the CH1 domains to be in close proximity of each other and displaces the chaperone BiP. With the addition of other chaperones in the ER such as a PPIase (orange) which isomerises a proline within the CH1 domain from a *trans* to a *cis* state and members of the PDI family (purple) to introduce the interchain disulphide between the CH1 domains, the completely folded HC dimer can then be secreted from the ER in a LC independent manner.

One approach to determine the mechanism behind how the CH1 domain is able to fold in the absence of LC would be to look at the crystal structure of the HC dimer. This information would allow one to observe how the CH1 and VH domains are interacting and the amino acids involved. Any amino acids identified could potentially reveal mutagenesis experiments to test for HC dimer secretion. A good evaluation of these findings would be to compare the crystal structure of the HC dimer with the crystal structure already solved for LC dimers and Fab fragments to note if the domains are interacting in the same way (Ely, Herron et al., 1989, Roussel et al., 1999, Saul & Poljak, 1992). The results in this thesis with the truncated HC A33 Fab show that this part of the protein is still able to fold, secrete and form an interchain disulphide between the two CH1 domains. This disulphide is probably also present in the full length HC dimer, which would indicate the overall predicted structure of the HC dimer would no longer represent the “Y” structure seen for IgG but would instead resemble a more “I” structure with the Fab HCs forming a dimer in the same way as the Fc fragment (Deisenhofer et al., 1976a, Deisenhofer et al., 1976b).

Although it has been shown that the CH1 domain within the HC dimer is folded, the conformation that it adopts in the quaternary structure has not been analysed. A partially folded CH1 domain or a CH1 domain that has adopted a different conformation may have occurred as the folding mechanism is different in the absence of LC. Without the information a crystal structure would provide, an alternative method to use to address this question could be to use computer modelling. Using the amino acid sequence and predicted quaternary structure, alternative structures can be modelled to hypothesise the final conformation of the CH1 domain within the HC dimer. For example this technology is already widely using in modelling the overall 3D structure of therapeutic antibodies using the amino acid sequence and previous knowledge on domain interactions and already solved crystal structures (Marcatili, Olimpieri et al., 2015).

It is known that the association of the VH and VL domains of the HC and LC is thought to be the initial domain interaction which envelops the complete folding of IgG via the displacement of BiP and action of the CL domain. However, if the VH domain has a stronger association for itself as a partner domain than the supposed VL domain, it can be hypothesised that the same initial domain interaction that causes IgG folding and secretion could be the same for HC dimer secretion. However, in the case of HC dimer secretion the VH domains are forming a complex with one another rather than the more well-known VL-VH interface. If this is the case, this action alone could initiate the folding of the CH1 domain and displace BiP, enabling the interaction of the two CH1 domains, which would not normally occur in IgG folding. As the CL domain displaces BiP by folding the CH1 domain but also spatially, the two CH1 domains interacting with one another in addition to being stabilised by the VH domain could explain the release of BiP in the absence of the CL domain. The formation of the interchain disulphide between the folded CH1 domains

would probably be the final stage of the folding of the CH1 domains as this is the case for the CH1 domain in the presence of the CL domain. The displacement of BiP from the CH1 domain during the process of HC dimer secretion would occur once the CH1 domains have folded enabling release of the HC dimer from the ER to be secreted from the cell. Engineering of variable domains which is seen for all therapeutic antibodies could result in unknown domain interactions shifting the equilibrium from a preferred VH-VL interface to a VH-VH or VL-VL interface which would result in the secretion of HC and LC dimers. Therefore, the development of therapeutic antibodies and variable domain engineering must be considered more carefully in terms of the potential effect of HC dimer secretion.

Therefore, from the initial results presented in this thesis, with the combination of sequence and structural techniques to pursue, understanding the mechanism behind HC dimer secretion is an achievable goal. This information could identify an unknown naturally occurring negative feedback mechanism in B cell development and contribute to the engineering of therapeutic antibodies to avoid the phenotype of HC dimer secretion.

6.4 Engineering of FabFvs to avoid multimer formation

A proposed model for the multimerisation of FabFvs was described in Chapter 5 which detailed the different intermediates formed in the FabFv assembly pathway. The hypothesis was that multimer formation of FabFvs occurs via the LC as when expressed alone it had the same phenotype as FabFvs of forming multimers which were secreted. It has also been suggested that the second VL domain of the LC could be orchestrating this multimerisation phenotype as the single LC does not yield further multimeric forms other than a dimer as seen in section 4.7.

In order to inhibit the formation of multimeric FabFvs, re-engineering could be investigated to improve the stability and organisation which could favour the monomeric form. The stability of the CH1 and CL interface of a Fab can be altered via mutating the interactions to increase the hydrophobic core which has been shown to increase the thermostability of the Fab fragment (Teerinen, Valjakka et al., 2006). This type of engineering could be investigated to shift the association of the HC and LCs rather than the formation of the HC dimer and LCHC₂ intermediate seen in the 2D gel electrophoresis. Furthermore, if the Fv portion of the FabFv was engineered to try to increase association and stability without the use of a disulphide, this could potentially reduce the formation of multimeric species of FabFvs. For example, in a scFv altering the orientation of the VH/VL assembly has shown to improve the level of expression and increase the association of the two domains (Ayyar, Hearty et al., 2015). The VH domain is said to be more stable than the VL domain and requires its intrachain disulphide to prevent aggregation and to promote the folding and association with the VL domain (Ramm, Gehrig et al., 1999).

During antibody folding and assembly, the initial interaction of the VH and VL domains drives association of the LC and HC. As the VH and VL domains of the Fv fragment interact and are required for function and structure, further engineering is needed to ensure that the interaction is with its partner chain and not with a VH domain from a separate FabFv HC. This manipulation will be difficult so as not to abolish the function of the Fv fragment. If this cannot be achieved; an alternative could be to investigate engineering the linker further to promote association between the two and inhibit the association with a different chain. For example, a helix-zipper was used on a scFv in place of CL and CH1 domains to inhibit multimerisation and aggregation which also increased its stability for use as a therapeutic (Arndt et al., 2001). If this helical association was introduced to the linker part of the FabFv then this could aid in stabilising the monomer and inhibit the formation of multimeric species.

6.5 Conclusions

One of the antibodies which have been extensively investigated in this thesis is A33 IgG4. This particular antibody has yielded several interesting and unexpected outcomes which have provided the main focus for the results which have been discussed here. For example, the LC of this antibody is unable to secrete dimers when transfected alone. However, the HC is cleaved within the cell and secreted as a dimer in the forms of full-length, truncated Fab and FabFv. However, the secretion of the FabFv dimer is not as abundant as seen with the other forms of the HC and this may be due to the additional VH domain present on the HC. Another consequence which is observed for the HC dimer secretion of the FabFv is the additional HC dimer band which has been hypothesised as a HC dimer which has not formed all its interchain disulphide bonds (one between the CH1 domains and one between the C-terminal VH domains). As the A33 Fab HC dimer only forms a single protein band suggesting only one disulphide bonded species exists, one can hypothesise that the additional VH domain with the engineered free thiol of the FabFv constructs is contributing to the two forms of HC dimer which are secreted by the A33 and 497 FabFvs. Therefore, if the additional C-terminal VH domain has or has not formed its interchain disulphide bond, this could attribute to the two different species which are secreted compared to the single species secreted for the A33 Fab. Furthermore, as HC dimer secretion is seen in the presence of LC for the full length antibody but not when the FabFv HC is co-expressed with its partner LC; the additional C-terminal variable domains could be inhibiting the process of HC dimer secretion. This inhibition could occur if this particular VH domain preferentially associates with its partner VL domain. If the association of the VH domain with the VL domain is stronger than the VH domain with itself, then HC dimer secretion may be inhibited.

As the secretion of HC dimers in the full length antibody has been shown to be VH domain specific, it would be useful to note whether these different variable domains retain their ability or lack of to secrete HC dimers in the FabFv form. All seven HC FabFvs were generated to investigate this hypothesis, however, unfortunately only A33 and 497 expressed well and sufficiently radiolabelled to observe the phenotype of HC dimer secretion. The reason for lack of expression of the other HC FabFvs could be that they require LC for expression, they are toxic to the cell, or the transfectability of the constructs is poor. This would be an area of troubleshooting which would need to be further addressed.

The variable domain of the antibody, specifically the VH domain, has been implicated in the ability or lack of to secrete HC dimers, but also has contributed to the cleavage of the A33 IgG4 HC. Furthermore, the C-terminal VL domain of the FabFvs investigated has been suggested to be the cause of multimer formation, and different VL domains have also been shown to contribute to the ability or lack of to secrete LC dimers (497 VL domain is able to secrete LC dimers and A33 VL domain is not). It is known that the main role of the variable domain is antigen binding; however, the results presented in this thesis suggest an additional role contributing to the folding and stability of the protein. The variable domain has been implicated in the ability of the LC or HC to form dimers and become secreted indicating it is playing a role in the folding of the overall protein. Also the presence of a cleavage product which is specific to the IgG4 isotype with the A33 VH domain suggests that the VH region may be also contributing to the overall stability of the protein and its susceptibility to cleavage. With this information, the engineering of the variable domain of therapeutic IgGs can be considered more carefully to the potential detrimental effects that may follow.

Chapter 7 : Appendices

7.1 Appendix 1

Table 3: Table of proteins produced after expression in CHO cells from constructs generated.

Construct	Protein(s) produced	Molecular weight as appears on denaturing gel (kDa)	Secreted Y/N
A33 LC	LC, LC ₂ , IgG	20, 40, 200	Y- small amount as LC dimer or as part of IgG
A33 HC (IgG1 and IgG4)	HC, HC ₂ , IgG	50, 100-135, 200	Y- as HC dimer or as part of IgG
A33 HC (P225S)	HC, HC ₂ , IgG	50, 100-135, 200	Y- as HC dimer or as part of IgG
A33 HC (IgG1 hinge)	HC, HC ₂ , IgG	50, 100-135, 200	Y- as HC dimer or as part of IgG
A33 HC (P151A)	HC, HC ₂ , IgG	50, 100-135, 200	N
A33 HC (humanised)	HC, HC ₂	50, 100-135	Y- as HC dimer to a lesser extent
A33 HC Fab	HC, HC ₂	25, 50	Y- as HC dimer
A33 HC Fab (C131S)	HC	25	Y
A33 HC FabFv	HC, HC ₂	30/35, 60	Y- as HC dimer
HC1	HC, HC ₂	50, 100-135	N
HC2	HC, HC ₂	50, 100-135	N
HC3	HC, HC ₂	50, 100-135	Y- as HC dimer
HC4	HC, HC ₂	50, 100-135	Y- as HC dimer
HC 5	HC, HC ₂	50, 100-135	N
497 LC	LC, LC ₂ , IgG	20, 40, 200	Y- as LC dimer
497 HC gH8	HC, HC ₂ , IgG	50, 100-135, 200	Y- as HC dimer

497 LC FabFv	LC, LC ₂ , LC ₃ , LCHC, LCHC ₂ , LCHC _n	30/35, 60, 80, 75, 175, 200+	Y- LC expressed alone secreted (LC, LC ₂ , LC ₃) -LC expressed with HC FabFv secreted (LCHC/ _n)
497 HC FabFv	HC, HC ₂ LCHC, LCHC ₂ , LCHC _n	30/40, 75 75, 175, 200+	Y- as HC dimer when expressed alone -HC expressed with LC FabFv secreted (LCHC/ _n)
A33 VH domain	Single immunoglobulin domain	13	N/A
CH1 domain	Single immunoglobulin domain	11	N/A
CL domain	Single immunoglobulin domain	13	N/A

7.2 Appendix 2

Table 4: Table of primers used for mutagenesis and sub-cloning protocols.

Primer	Oligonucleotide
A33 Nde1 (F)	GCCGCCCATATGGAAGTTAAACTGG
A33 Hind3 (R)	CCGGCGAAGCTTTTACGATGACACCGTGACCAGGGTGCCTTGACCCC
CH1 Nde1 (F)	CCCGCCCATATGGCGAGTACCAAAGG
CH1 Hind3 (R)	CCCGGCAAGCTTTTAACCATATTTACTTTCAACACGTTTATCCACTTTCGTATTACTC GG
CL Nde1 (F)	GCCGCCCATATGGCGGCCCGAGCG
CL Hind3 (R)	GGCGGGAAGCTTTTATTCACCGCGGTTGAAGGATTTCTGAACCGGGGAACCTCAGG CCTTGG
A33 Fab (F)	CCTTGACACGAAGCTTGCCACCATGG
A33 Fab (R)	GGGCGGCCGCTTAAGTGTCTAGACCCAATAGCGGATTGGGTATTGGTTTGCCCTC ACCAACTCTCTTGTCC
C131S (F)	CCATCCGTCTTCCCCCTGGCGCCCAGCTCCAGGAGCACCTCCGAGAGCACAGCC
C131S (R)	GGCTGTGCTCTCGGAGGTGCTCCTGGAGCTGGGCGCCAGGGGGAAGACGGATGG
P151A (F)	GCCTGGTCAAGGACTACTTCGCCGAACCGGTGACGGTGTCGTGG
P151A (R)	CCACGACACCGTCACCGGTTGCGCGAAGTAGTCCTTGACCAGGC
P228S (F)	CCAAATATGGTCCCCCATGCCATCATGCCAGGTAAGCCAACCCAGG
P228S (R)	CCTGGGTTGGCTTACCTGGGCATGATGGGCATGGGGGACCATATTTGG

7.3 Appendix 3

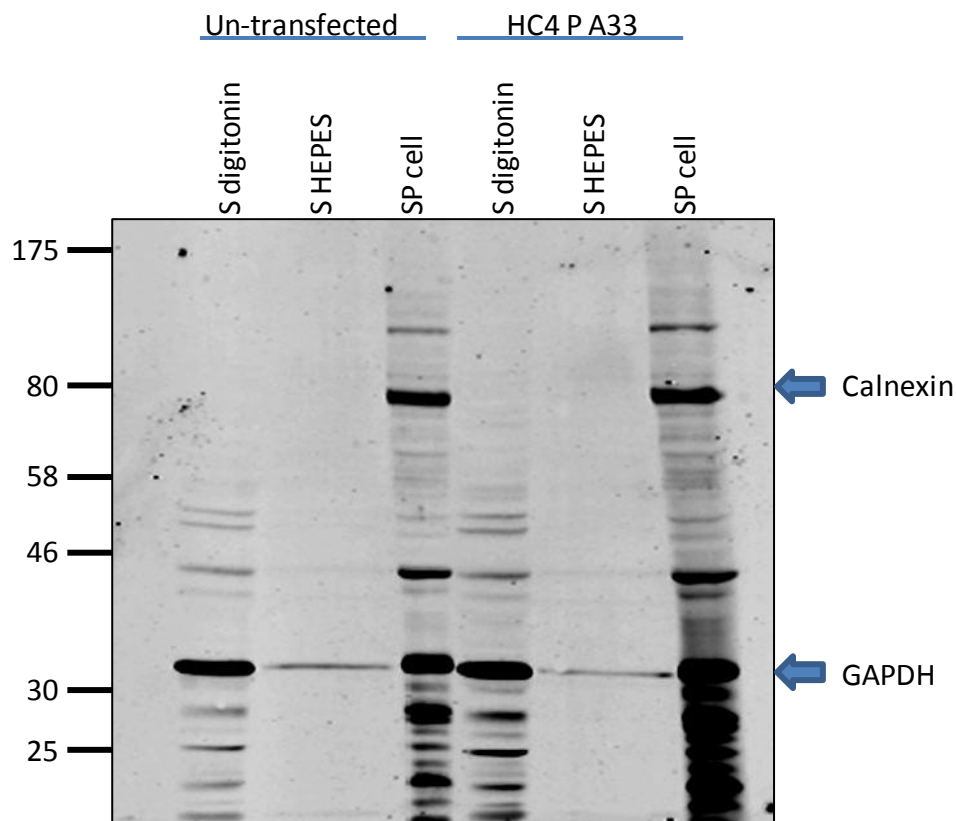


Figure 7.1: Appendix 2: Confirmation of separation of the endoplasmic reticulum from the cytosol of the cell.

Each sample from the SP cell prep to fractionate the compartments of the cell were run on a gradient gel (4-20%) (Chapter 3 Figure 3.3E). Two proteins were looked at to identify whether or not the ER was successfully separated from the cytosol. GAPDH was used as it is a protein which resides in the cytosol of the cell, and Calnexin was used as this is a protein which exists solely in the ER of the cell. From the western blot above, it can be seen that Calnexin is present in both SP cell preps and not in the digitonin or HEPES fractions confirming successful isolation of the ER. GAPDH is present in all three samples, mainly the digitonin and SP cell fractions. This confirms the digitonin fraction is a representative of the cytosol however some cytosolic proteins are present in the SP cell prep as this is never 100% effective. However as the cleavage product is not seen in the digitonin fraction and GAPDH is, then it can be said that it resides in the ER of the cell.

7.4 Appendix 4

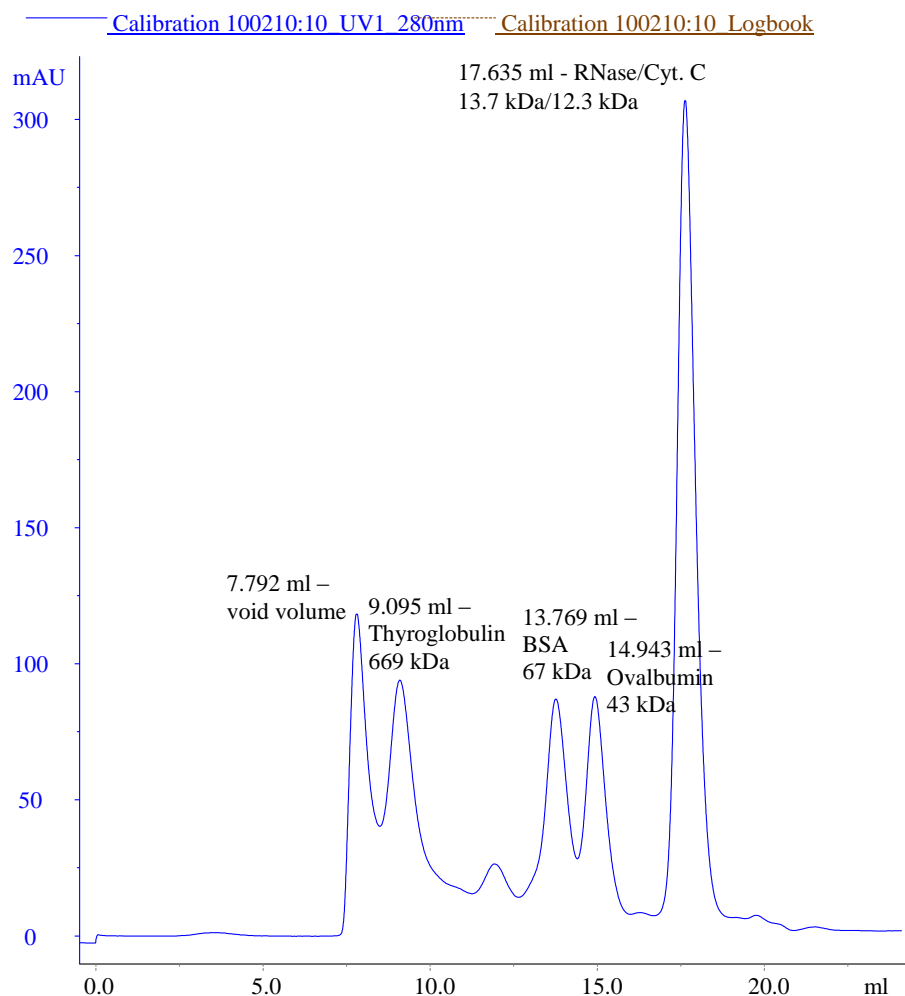


Figure 7.2 Calibration data using protein standards for Superdex 200 10 300 gel filtration column.

Figure and calibrations were provided by Dr. Ojore Oka. The calibration of the Superdex 200 10 30 gel filtration column is able to elute a range of protein standards with molecular weights ranging from 669 kDa to 12.3 kDa which is suitable for the molecular weight of the individual immunoglobulin domains which require purification.

List of References

- Aalberse RC, Schuurman J (2002) IgG4 breaking the rules. *Immunology* 105: 9-19
- Ailor E, Betenbaugh MJ (1998) Overexpression of a cytosolic chaperone to improve solubility and secretion of a recombinant IgG protein in insect cells. *Biotechnol Bioeng* 58: 196-203
- An Y, Zhang Y, Mueller HM, Shameem M, Chen X (2014) A new tool for monoclonal antibody analysis: application of IdeS proteolysis in IgG domain-specific characterization. *MAbs* 6: 879-93
- Angal S, King DJ, Bodmer MW, Turner A, Lawson AD, Roberts G, Pedley B, Adair JR (1993) A single amino acid substitution abolishes the heterogeneity of chimeric mouse/human (IgG4) antibody. *Molecular immunology* 30: 105-8
- Arndt KM, Müller KM, Plückthun A (2001) Helix-stabilized Fv (hsFv) antibody fragments: substituting the constant domains of a Fab fragment for a heterodimeric coiled-coil domain. *J Mol Biol* 312: 221-8
- Ayyar BV, Hearty S, O'Kennedy R (2015) Facile domain rearrangement abrogates expression recalcitrance in a rabbit scFv. *Appl Microbiol Biotechnol* 99: 2693-703
- Baselga J, Tripathy D, Mendelsohn J, Baughman S, Benz CC, Dantis L, Sklarin NT, Seidman AD, Hudis CA, Moore J, Rosen PP, Twaddell T, Henderson IC, Norton L (1996) Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. *J Clin Oncol* 14: 737-44
- Bause E (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *The Biochemical journal* 209: 331-6
- Berasain P, Carmona C, Frangione B, Cazzulo JJ, Goñi F (2003) Specific cleavage sites on human IgG subclasses by cruzipain, the major cysteine proteinase from *Trypanosoma cruzi*. *Mol Biochem Parasitol* 130: 23-9
- Berasain P, Carmona C, Frangione B, Dalton JP, Goñi F (2000) *Fasciola hepatica*: parasite-secreted proteinases degrade all human IgG subclasses: determination of the specific cleavage sites and identification of the immunoglobulin fragments produced. *Exp Parasitol* 94: 99-110
- Berg JM, Tymoczko JL, Stryer L (2006) *Biochemistry* sixth edition. New York
- Bergman LW, Kuehl WM (1979) Formation of an intrachain disulfide bond on nascent immunoglobulin light chains. *J Biol Chem* 254: 8869-76
- Blackstone EA, Joseph PF (2013) The Economics of Biosimilars. *American Health and Drug Benefits* 8: 469-478
- Bole DG, Hendershot LM, Kearney JF (1986) Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. *The Journal of cell biology* 102: 1558-66

- Boulianne GL, Hozumi N, Shulman MJ (1984) Production of functional chimaeric mouse/human antibody. *Nature* 312: 643-6
- Braakman I, Bulleid NJ (2011) Protein folding and modification in the mammalian endoplasmic reticulum. *Annual review of biochemistry* 80: 71-99
- Brezski RJ, Jordan RE (2010) Cleavage of IgGs by proteases associated with invasive diseases: an evasion tactic against host immunity? *MAbs* 2: 212-20
- Cain K, Peters S, Hailu H, Sweeney B, Stephens P, Heads J, Sarkar K, Ventom A, Page C, Dickson A (2013) A CHO cell line engineered to express XBP1 and ERO1- α has increased levels of transient protein expression. *Biotechnology progress* 29: 697-706
- Campion SR (2016) Conserved aromatic residues as determinants in the folding and assembly of immunoglobulin variable domains. *Molecular immunology* 70: 63-71
- Carter PJ (2006) Potent antibody therapeutics by design. *Nature reviews Immunology* 6: 343-57
- Chakrabarti A, Chen AW, Varner JD (2011) A review of the mammalian unfolded protein response. *Biotechnology and bioengineering* 108: 2777-93
- Chan AC, Carter PJ (2010) Therapeutic antibodies for autoimmunity and inflammation. *Nature reviews Immunology* 10: 301-16
- Chothia C, Novotný J, Brucoleri R, Karplus M (1985) Domain association in immunoglobulin molecules. The packing of variable domains. *Journal of molecular biology* 186: 651-63
- Clark M (2000) Antibody humanization: a case of the 'Emperor's new clothes'? *Immunol Today* 21: 397-402
- Clynes RA, Towers TL, Presta LG, Ravetch JV (2000) Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. *Nat Med* 6: 443-6
- Cordoba AJ, Shyong BJ, Breen D, Harris RJ (2005) Non-enzymatic hinge region fragmentation of antibodies in solution. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* 818: 115-21
- Correia IR (2010) Stability of IgG isotypes in serum. *MAbs* 2: 221-32
- Costa TE, Franke RR, Sanchez M, Misulovin Z, Nussenzweig MC (1992) Functional reconstitution of an immunoglobulin antigen receptor in T cells. *J Exp Med* 175: 1669-76
- Davies J, Riechmann L (1996) Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. *Protein engineering* 9: 531-7

Davis DP, Khurana R, Meredith S, Stevens FJ, Argon Y (1999) Mapping the major interaction between binding protein and Ig light chains to sites within the variable domain. *J Immunol* 163: 3842-50

de Haard HJ, Kazemier B, van der Bent A, Oudshoorn P, Boender P, van Gemen B, Arends JW, Hoogenboom HR (1998) Absolute conservation of residue 6 of immunoglobulin heavy chain variable regions of class IIA is required for correct folding. *Protein Eng* 11: 1267-76

de StGroth SF, Scheidegger D (1980) Production of monoclonal antibodies: strategy and tactics. *J Immunol Methods* 35: 1-21

Deisenhofer J, Colman PM, Epp O, Huber R (1976a) Crystallographic structural studies of a human Fc fragment. II. A complete model based on a Fourier map at 3.5 Å resolution. *Hoppe Seylers Z Physiol Chem* 357: 1421-34

Deisenhofer J, Colman PM, Huber R, Haupt H, Schwick G (1976b) Crystallographic structural studies of a human Fc-fragment. I. An electron-density map at 4 Å resolution and a partial model. *Hoppe Seylers Z Physiol Chem* 357: 435-45

Dunkelberger JR, Song WC (2010a) Complement and its role in innate and adaptive immune responses. *Cell Res* 20: 34-50

Dunkelberger JR, Song WC (2010b) Role and mechanism of action of complement in regulating T cell immunity. *Mol Immunol* 47: 2176-86

Durocher Y, Perret S, Kamen A (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res* 30: E9

Elkabetz Y, Argon Y, Bar-Nun S (2005) Cysteines in CH1 underlie retention of unassembled Ig heavy chains. *J Biol Chem* 280: 14402-12

Elliott JG, Oliver JD, High S (1997) The thiol-dependent reductase ERp57 interacts specifically with N-glycosylated integral membrane proteins. *The Journal of biological chemistry* 272: 13849-55

Ely KR, Herron JN, Harker M, Edmundson AB (1989) Three-dimensional structure of a light chain dimer crystallized in water. Conformational flexibility of a molecule in two crystal forms. *J Mol Biol* 210: 601-15

Ewert S, Honegger A, Plückthun A (2004) Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. *Methods* 34: 184-99

Ewert S, Huber T, Honegger A, Plückthun A (2003) Biophysical properties of human antibody variable domains. *Journal of molecular biology* 325: 531-53

Feige MJ, Groscurth S, Marciniowski M, Shimizu Y, Kessler H, Hendershot LM, Buchner J (2009) An unfolded CH1 domain controls the assembly and secretion of IgG antibodies. *Molecular cell* 34: 569-79

- Feige MJ, Hagn F, Esser J, Kessler H, Buchner J (2007) Influence of the internal disulfide bridge on the folding pathway of the CL antibody domain. *Journal of molecular biology* 365: 1232-44
- Feige MJ, Hendershot LM, Buchner J (2010) How antibodies fold. *Trends in biochemical sciences* 35: 189-98
- Feige MJ, Walter S, Buchner J (2004) Folding mechanism of the CH2 antibody domain. *Journal of molecular biology* 344: 107-18
- Ferrante A, Beard LJ, Feldman RG (1990) IgG subclass distribution of antibodies to bacterial and viral antigens. *Pediatr Infect Dis J* 9: S16-24
- Freeze HH, Kranz C (2010) Endoglycosidase and glycoamidase release of N-linked glycans. *Curr Protoc Protein Sci Chapter 12: Unit12.4*
- García-Gómez E, Vaca S, Pérez-Méndez A, Ibarra-Caballero J, Pérez-Márquez V, Tenorio VR, Negrete-Abascal E (2005) Gallibacterium anatis-secreted metalloproteases degrade chicken IgG. *Avian Pathol* 34: 426-9
- Gearing AJ, Thorpe SJ, Miller K, Mangan M, Varley PG, Dudgeon T, Ward G, Turner C, Thorpe R (2002) Selective cleavage of human IgG by the matrix metalloproteinases, matrilysin and stromelysin. *Immunol Lett* 81: 41-8
- Grant SR, Welply JK, Olson EN, Lennarz WJ (1986) Oligosaccharyltransferase activity is markedly increased during differentiation of a nonfusing myoblast cell line. *Archives of biochemistry and biophysics* 248: 424-8
- Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF (1995) A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* 374: 168-73
- Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* 1: 2876-90
- Grey HM, Kunkel HG (1964) H Chain Subgroups of myeloma proteins and normal 7S gamma-globulin. *The Journal of experimental medicine* 120: 253-66
- Halaby DM, Poupon A, Moron J (1999) The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng* 12: 563-71
- Hamel PA, Isenman DE, Klein MH, Luedtke R, Dorrington KJ (1984) Structural basis for the preferential association of autologous immunoglobulin subunits: role of the J region of the light chain. *Molecular immunology* 21: 277-83
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R (1993) Naturally occurring antibodies devoid of light chains. *Nature* 363: 446-8

Harmsen MM, van Solt CB, van Zijderveld-van Bemmelen AM, Niewold TA, van Zijderveld FG (2006) Selection and optimization of proteolytically stable llama single-domain antibody fragments for oral immunotherapy. *Appl Microbiol Biotechnol* 72: 544-51

Harris LJ, Larson SB, Hasel KW, McPherson A (1997) Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* 36: 1581-97

Hatahet F, Ruddock LW (2007) Substrate recognition by the protein disulfide isomerases. *The FEBS journal* 274: 5223-34

Heads JT, Adams R, D'Hooghe LE, Page MJ, Humphreys DP, Popplewell AG, Lawson AD, Henry AJ (2012) Relative stabilities of IgG1 and IgG4 Fab domains: influence of the light-heavy interchain disulfide bond architecture. *Protein science : a publication of the Protein Society* 21: 1315-22

Heath JK, White SJ, Johnstone CN, Catimel B, Simpson RJ, Moritz RL, Tu GF, Ji H, Whitehead RH, Groenen LC, Scott AM, Ritter G, Cohen L, Welt S, Old LJ, Nice EC, Burgess AW (1997) The human A33 antigen is a transmembrane glycoprotein and a novel member of the immunoglobulin superfamily. *Proceedings of the National Academy of Sciences of the United States of America* 94: 469-74

Helenius A, Aebi M (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 73: 1019-49

Hellman R, Vanhove M, Lejeune A, Stevens FJ, Hendershot LM (1999) The in vivo association of BiP with newly synthesized proteins is dependent on the rate and stability of folding and not simply on the presence of sequences that can bind to BiP. *J Cell Biol* 144: 21-30

Hendershot L, Bole D, Köhler G, Kearney JF (1987) Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. *J Cell Biol* 104: 761-7

Ho SC, Wang T, Song Z, Yang Y (2015) IgG Aggregation Mechanism for CHO Cell Lines Expressing Excess Heavy Chains. *Molecular biotechnology* 57: 625-34

Hochman J, Gavish M, Inbar D, Givol D (1976) Folding and interaction of subunits at the antibody combining site. *Biochemistry* 15: 2706-10

Hoebe K, Janssen E, Beutler B (2004) The interface between innate and adaptive immunity. *Nat Immunol* 5: 971-4

Hsu TA, Eiden JJ, Bourgarel P, Meo T, Betenbaugh MJ (1994) Effects of co-expressing chaperone BiP on functional antibody production in the baculovirus system. *Protein Expr Purif* 5: 595-603

Hsu TA, Watson S, Eiden JJ, Betenbaugh MJ (1996) Rescue of immunoglobulins from insolubility is facilitated by PDI in the baculovirus expression system. *Protein Expr Purif* 7: 281-8

Hwang WY, Almagro JC, Buss TN, Tan P, Foote J (2005) Use of human germline genes in a CDR homology-based approach to antibody humanization. *Methods* 36: 35-42

Igawa T, Tsunoda H, Kikuchi Y, Yoshida M, Tanaka M, Koga A, Sekimori Y, Orita T, Aso Y, Hattori K, Tsuchiya M (2010) VH/VL interface engineering to promote selective expression and inhibit conformational isomerization of thrombopoietin receptor agonist single-chain diabody. *Protein Eng Des Sel* 23: 667-77

Ishii Y, Murakami J, Sasaki K, Tsukahara M, Wakamatsu K (2014) Efficient folding/assembly in Chinese hamster ovary cells is critical for high quality (low aggregate content) of secreted trastuzumab as well as for high production: stepwise multivariate regression analyses. *Journal of bioscience and bioengineering* 118: 223-30

Jakob CG, Edalji R, Judge RA, DiGiammarino E, Li Y, Gu J, Ghayur T (2013) Structure reveals function of the dual variable domain immunoglobulin (DVD-Ig™) molecule. *MAbs* 5: 358-63

Jayapal KP, Wlaschin KF, Hu W-S, Yap MGS (2007) Recombinant protein therapeutics from CHO cells-20 years and counting. *Chemical engineering progress* 103: 40-47

Jefferis R, Weston PD, Stanworth DR, Clamp JR (1968) Relationship between the papain sensitivity of human gammaG immunoglobulins and their heavy chain subclass. *Nature* 219: 646-9

Jessop CE, Watkins RH, Simmons JJ, Tasab M, Bulleid NJ (2009) Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins. *Journal of cell science* 122: 4287-95

Johansen FE, Braathen R, Brandtzaeg P (2000) Role of J chain in secretory immunoglobulin formation. *Scand J Immunol* 52: 240-8

Jones PT, Dear PH, Foote J, Neuberger MS, Winter G (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* 321: 522-5

Jutel M, Akdis CA (2011) Immunological mechanisms of allergen-specific immunotherapy. *Allergy* 66: 725-32

Kaern M, Blake WJ, Collins JJ (2003) The engineering of gene regulatory networks. *Annual review of biomedical engineering* 5: 179-206

Kats M, Richberg PC, Hughes DE (1995) Conformational diversity and conformational transitions of a monoclonal antibody monitored by circular dichroism and capillary electrophoresis. *Anal Chem* 67: 2943-8

Kim EJ, Kim BS, Choi DB, Chi SG, Choi TH (2016) Enhanced tumor retention of radioiodinated anti-epidermal growth factor receptor antibody using novel bifunctional iodination linker for radioimmunotherapy. *Oncology reports*

Kirkpatrick RB, Ganguly S, Angelichio M, Griego S, Shatzman A, Silverman C, Rosenberg M (1995) Heavy chain dimers as well as complete antibodies are efficiently formed and secreted from *Drosophila* via a BiP-mediated pathway. *J Biol Chem* 270: 19800-5

- Knappik A, Ge L, Honegger A, Pack P, Fischer M, Wellnhofer G, Hoess A, Wölle J, Plückthun A, Virnekäs B (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* 296: 57-86
- Knittler MR, Haas IG (1992) Interaction of BiP with newly synthesized immunoglobulin light chain molecules: cycles of sequential binding and release. *The EMBO journal* 11: 1573-81
- Krambeck FJ, Bennun SV, Narang S, Choi S, Yarema KJ, Betenbaugh MJ (2009) A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology* 19: 1163-75
- Ku SC, Ng DT, Yap MG, Chao SH (2008) Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. *Biotechnology and bioengineering* 99: 155-64
- Köhler G, Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256: 495-7
- Lazar GA, Desjarlais JR, Jacinto J, Karki S, Hammond PW (2007) A molecular immunology approach to antibody humanization and functional optimization. *Molecular immunology* 44: 1986-98
- Lee YK, Brewer JW, Hellman R, Hendershot LM (1999) BiP and immunoglobulin light chain cooperate to control the folding of heavy chain and ensure the fidelity of immunoglobulin assembly. *Mol Biol Cell* 10: 2209-19
- Leitzgen K, Knittler MR, Haas IG (1997) Assembly of immunoglobulin light chains as a prerequisite for secretion. A model for oligomerization-dependent subunit folding. *J Biol Chem* 272: 3117-23
- Lewis Phillips GD, Li G, Dugger DL, Crocker LM, Parsons KL, Mai E, Blättler WA, Lambert JM, Chari RV, Lutz RJ, Wong WL, Jacobson FS, Koeppen H, Schwall RH, Kenkare-Mitra SR, Spencer SD, Sliwkowski MX (2008) Targeting HER2-positive breast cancer with trastuzumab-DM1, an antibody-cytotoxic drug conjugate. *Cancer Res* 68: 9280-90
- Lewis SM, Wu X, Pustilnik A, Sereno A, Huang F, Rick HL, Guntas G, Leaver-Fay A, Smith EM, Ho C, Hansen-Estruch C, Chamberlain AK, Truhlar SM, Conner EM, Atwell S, Kuhlman B, Demarest SJ (2014) Generation of bispecific IgG antibodies by structure-based design of an orthogonal Fab interface. *Nature biotechnology* 32: 191-8
- Lilie H, Rudolph R, Buchner J (1995) Association of antibody chains at different stages of folding: prolyl isomerization occurs after formation of quaternary structure. *J Mol Biol* 248: 190-201
- Lipson C, Alalouf G, Bajorek M, Rabinovich E, Atir-Lande A, Glickman M, Bar-Nun S (2008) A proteasomal ATPase contributes to dislocation of endoplasmic reticulum-associated degradation (ERAD) substrates. *The Journal of biological chemistry* 283: 7166-75
- Liu H, Bulseco GG, Sun J (2006) Effect of posttranslational modifications on the thermal stability of a recombinant monoclonal antibody. *Immunol Lett* 106: 144-53

- Liu H, May K (2012) Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. *mAbs* 4: 17-23
- Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J (2000) *Molecular Cell Biology*. 4th edition. W. H. Freeman, New York
- Lu D, Jimenez X, Zhang H, Bohlen P, Witte L, Zhu Z (2002) Fab-scFv fusion protein: an efficient approach to production of bispecific antibody fragments. *Journal of immunological methods* 267: 213-26
- Mains PE, Sibley CH (1983) The requirement of light chain for the surface deposition of the heavy chain of immunoglobulin M. *The Journal of biological chemistry* 258: 5027-33
- Marcatili P, Olimpieri PP, Chailyan A, Tramontano A (2015) Erratum: antibody modeling using the Prediction of ImmunoGlobulin Structure (PIGS) web server. *Nature protocols* 10: 644
- Marcel V, Perrier S, Aoubala M, Ageorges S, Groves MJ, Diot A, Fernandes K, Tauro S, Bourdon JC (2010) $\Delta 160p53$ is a novel N-terminal p53 isoform encoded by $\Delta 133p53$ transcript. *FEBS Lett* 584: 4463-8
- Martsev SP, Dubnovitsky AP, Stremovsky OA, Chumanevich AA, Tsybovsky YI, Kravchuk ZI, Deyev SM (2002) Partially structured state of the functional VH domain of the mouse anti-ferritin antibody F11. *FEBS letters* 518: 177-82
- Mason M, Sweeney B, Cain K, Stephens P, Sharfstein ST (2012) Identifying bottlenecks in transient and stable production of recombinant monoclonal-antibody sequence variants in Chinese hamster ovary cells. *Biotechnology progress* 28: 846-55
- May MB, Glode A (2016) Blinatumomab: A novel, bispecific, T-cell engaging antibody. *Am J Health Syst Pharm* 73: e6-e13
- Milenic DE, Brady ED, Brechbiel MW (2004) Antibody-targeted radiation cancer therapy. *Nat Rev Drug Discov* 3: 488-99
- Minegishi Y, Conley ME (2001) Negative selection at the pre-BCR checkpoint elicited by human mu heavy chains with unusual CDR3 regions. *Immunity* 14: 631-41
- Mitić N, Miraula M, Selleck C, Hadler KS, Uribe E, Pedroso MM, Schenk G (2014) Catalytic mechanisms of metallohydrolases containing two metal ions. *Adv Protein Chem Struct Biol* 97: 49-81
- Molla A, Kagimoto T, Maeda H (1988) Cleavage of immunoglobulin G (IgG) and IgA around the hinge region by proteases from *Serratia marcescens*. *Infect Immun* 56: 916-20
- Morrison SL, Johnson MJ, Herzenberg LA, Oi VT (1984) Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proceedings of the National Academy of Sciences of the United States of America* 81: 6851-5

- Morrison SL, Scharff MD (1975) Heavy chain-producing variants of a mouse myeloma cell line. *J Immunol* 114: 655-9
- Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezcan I, Sanal O, Bertrand Y, Philippe N, Fischer A, de Villartay JP (2001) Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* 105: 177-86
- Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA, Hamers R (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein engineering* 7: 1129-35
- Mårtensson IL, Ceredig R (2000) Review article: role of the surrogate light chain and the pre-B-cell receptor in mouse B-cell development. *Immunology* 101: 435-41
- Nisonoff A, Rivers MM (1961) Recombination of a mixture of univalent antibody fragments of different specificity. *Arch Biochem Biophys* 93: 460-2
- Oka OB, Yeoh HY, Bulleid NJ (2015) Thiol-disulfide exchange between the PDI family of oxidoreductases negates the requirement for an oxidase or reductase for each enzyme. *The Biochemical journal* 469: 279-88
- Ozato K, Mayer N, Sachs DH (1980) Hybridoma cell lines secreting monoclonal antibodies to mouse H-2 and Ia antigens. *J Immunol* 124: 533-40
- Parekh BS, Berger E, Sibley S, Cahya S, Xiao L, LaCerte MA, Vaillancourt P, Wooden S, Gately D (2012) Development and validation of an antibody-dependent cell-mediated cytotoxicity-reporter gene assay. *mAbs* 4: 310-8
- Peng SL, Szabo SJ, Glimcher LH (2002) T-bet regulates IgG class switching and pathogenic autoantibody production. *Proc Natl Acad Sci U S A* 99: 5545-50
- Peters SJ, Smales CM, Henry AJ, Stephens PE, West S, Humphreys DP (2012) Engineering an improved IgG4 molecule with reduced disulfide bond heterogeneity and increased Fab domain thermal stability. *The Journal of biological chemistry* 287: 24525-33
- Phalipon A, Cardona A, Kraehenbuhl JP, Edelman L, Sansonetti PJ, Corthésy B (2002) Secretory component: a new role in secretory IgA-mediated immune exclusion in vivo. *Immunity* 17: 107-15
- Pieper K, Grimbacher B, Eibel H (2013) B-cell biology and development. *J Allergy Clin Immunol* 131: 959-71
- Pybus LP, Dean G, West NR, Smith A, Daramola O, Field R, Wilkinson SJ, James DC (2014) Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and bioengineering* 111: 372-85
- Qi J, Ye X, Ren G, Kan F, Zhang Y, Guo M, Zhang Z, Li D (2014) Pharmacological efficacy of anti-IL-1 β scFv, Fab and full-length antibodies in treatment of rheumatoid arthritis. *Molecular immunology* 57: 59-65

- Raju TS, Scallon BJ (2006) Glycosylation in the Fc domain of IgG increases resistance to proteolytic cleavage by papain. *Biochemical and biophysical research communications* 341: 797-803
- Ramm K, Gehrig P, Plückthun A (1999) Removal of the conserved disulfide bridges from the scFv fragment of an antibody: effects on folding kinetics and aggregation. *J Mol Biol* 290: 535-46
- Raso V, Ritz J, Basala M, Schlossman SF (1982) Monoclonal antibody-ricin A chain conjugate selectively cytotoxic for cells bearing the common acute lymphoblastic leukemia antigen. *Cancer Res* 42: 457-64
- Ridgway JB, Presta LG, Carter P (1996) 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng* 9: 617-21
- Rita Costa A, Elisa Rodrigues M, Henriques M, Azeredo J, Oliveira R (2010) Guidelines to cell engineering for monoclonal antibody production. *Eur J Pharm Biopharm* 74: 127-38
- Ritter C, Helenius A (2000) Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nature structural biology* 7: 278-80
- Roberts S, Cheetham JC, Rees AR (1987) Generation of an antibody with enhanced affinity and specificity for its antigen by protein engineering. *Nature* 328: 731-4
- Rodrigues ME, Costa AR, Henriques M, Azeredo J, Oliveira R (2010) Technological progresses in monoclonal antibody production systems. *Biotechnology progress* 26: 332-51
- Roitt I, Brostoff J, Male D (1998) *Immunology Fifth Edition*. Butler and Tanner, Somerset, UK
- Rothlisberger D, Honegger A, Pluckthun A (2005) Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *Journal of molecular biology* 347: 773-89
- Rousseaux J, Aubert JP, Loucheux-Lefebvre MH (1982) Comparative study of the conformational features of rat immunoglobulin G subclasses by circular dichroism. *Biochim Biophys Acta* 701: 93-101
- Roussel A, Spinelli S, Déret S, Navaza J, Aucoeur P, Cambillau C (1999) The structure of an entire noncovalent immunoglobulin kappa light-chain dimer (Bence-Jones protein) reveals a weak and unusual constant domains association. *Eur J Biochem* 260: 192-9
- Roux KH, Strelets L, Michaelsen TE (1997) Flexibility of human IgG subclasses. *J Immunol* 159: 3372-82
- Röttger S, White J, Wandall HH, Olivo JC, Stark A, Bennett EP, Whitehouse C, Berger EG, Clausen H, Nilsson T (1998) Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *J Cell Sci* 111 (Pt 1): 45-60

- Sandberg H, Lutkemeyer D, Kuprin S, Wrangel M, Almstedt A, Persson P, Ek V, Mikaelsson M (2006) Mapping and partial characterization of proteases expressed by a CHO production cell line. *Biotechnol Bioeng* 95: 961-71
- Sarma JV, Ward PA (2011) The complement system. *Cell Tissue Res* 343: 227-35
- Saul FA, Poljak RJ (1992) Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 Å resolution. *Proteins* 14: 363-71
- Schaefer G, Haber L, Crocker LM, Shia S, Shao L, Dowbenko D, Totpal K, Wong A, Lee CV, Stawicki S, Clark R, Fields C, Lewis Phillips GD, Prell RA, Danilenko DM, Franke Y, Stephan JP, Hwang J, Wu Y, Bostrom J et al. (2011) A two-in-one antibody against HER3 and EGFR has superior inhibitory activity compared with monospecific antibodies. *Cancer Cell* 20: 472-86
- Schatz DG, Ji Y (2011) Recombination centres and the orchestration of V(D)J recombination. *Nature reviews Immunology* 11: 251-63
- Schellenberger EA, Weissleder R, Josephson L (2004) Optimal modification of annexin V with fluorescent dyes. *Chembiochem* 5: 271-4
- Schroeder HW, Cavacini L (2010) Structure and function of immunoglobulins. *J Allergy Clin Immunol* 125: S41-52
- Schuurman J, Van Ree R, Perdok GJ, Van Doorn HR, Tan KY, Aalberse RC (1999) Normal human immunoglobulin G4 is bispecific: it has two different antigen-combining sites. *Immunology* 97: 693-8
- Scott AM, Wolchok JD, Old LJ (2012) Antibody therapy of cancer. *Nat Rev Cancer* 12: 278-87
- Shaffer AL, Schlissel MS (1997) A truncated heavy chain protein relieves the requirement for surrogate light chains in early B cell development. *J Immunol* 159: 1265-75
- Siber GR, Schur PH, Aisenberg AC, Weitzman SA, Schiffman G (1980) Correlation between serum IgG-2 concentrations and the antibody response to bacterial polysaccharide antigens. *N Engl J Med* 303: 178-82
- Simpson ER, Herold EM, Buchner J (2009) The folding pathway of the antibody V(L) domain. *Journal of molecular biology* 392: 1326-38
- Sjodahl J (1977) Repetitive sequences in protein A from *Staphylococcus aureus*. Arrangement of five regions within the protein, four being highly homologous and Fc-binding. *Eur J Biochem* 73: 343-51
- Sjödahl J (1977) Structural studies on the four repetitive Fc-binding regions in protein A from *Staphylococcus aureus*. *Eur J Biochem* 78: 471-90

Skowronek MH, Hendershot LM, Haas IG (1998) The variable domain of nonassembled Ig light chains determines both their half-life and binding to the chaperone BiP. *Proceedings of the National Academy of Sciences of the United States of America* 95: 1574-8

Slanina H, Mündlein S, Hebling S, Schubert-Unkmeir A (2014) Role of epidermal growth factor receptor signaling in the interaction of *Neisseria meningitidis* with endothelial cells. *Infect Immun* 82: 1243-55

Spiess C, Zhai Q, Carter PJ (2015) Alternative molecular formats and therapeutic applications for bispecific antibodies. *Mol Immunol* 67: 95-106

Strop P, Ho WH, Boustany LM, Abdiche YN, Lindquist KC, Farias SE, Rickert M, Appah CT, Pascua E, Radcliffe T, Sutton J, Chaparro-Riggers J, Chen W, Casas MG, Chin SM, Wong OK, Liu SH, Vergara G, Shelton D, Rajpal A et al. (2012) Generating bispecific human IgG1 and IgG2 antibodies from any antibody pair. *J Mol Biol* 420: 204-19

Tan PH, Sandmaier BM, Stayton PS (1998) Contributions of a highly conserved VH/VL hydrogen bonding interaction to scFv folding stability and refolding efficiency. *Biophys J* 75: 1473-82

Tavender TJ, Bulleid NJ (2010) Molecular mechanisms regulating oxidative activity of the Ero1 family in the endoplasmic reticulum. *Antioxidants & redox signaling* 13: 1177-87

Teerinen T, Valjakka J, Rouvinen J, Takkinen K (2006) Structure-based stability engineering of the mouse IgG1 Fab fragment by modifying constant domains. *J Mol Biol* 361: 687-97

Thies MJ, Mayer J, Augustine JG, Frederick CA, Lilie H, Buchner J (1999) Folding and association of the antibody domain CH3: prolyl isomerization precedes dimerization. *Journal of molecular biology* 293: 67-79

Thies MJ, Talamo F, Mayer M, Bell S, Ruoppolo M, Marino G, Buchner J (2002) Folding and oxidation of the antibody domain C(H)3. *Journal of molecular biology* 319: 1267-77

Tijink BM, Laeremans T, Budde M, Stigter-van Walsum M, Dreier T, de Haard HJ, Leemans CR, van Dongen GA (2008) Improved tumor targeting of anti-epidermal growth factor receptor Nanobodies through albumin binding: taking advantage of modular Nanobody technology. *Mol Cancer Ther* 7: 2288-97

Tsybovsky Y, Shubenok DV, Kravchuk ZI, Martsev SP (2007) Folding of an antibody variable domain in two functional conformations in vitro: calorimetric and spectroscopic study of the anti-ferritin antibody VL domain. *Protein Eng Des Sel* 20: 481-90

Turner MW, Bennich HH, Natvig JB (1970) Pepsin digestion of human G-myeloma proteins of different subclasses. I. The characteristic features of pepsin cleavage as a function of time. *Clin Exp Immunol* 7: 603-25

Vanhove M, Usherwood YK, Hendershot LM (2001) Unassembled Ig heavy chains do not cycle from BiP in vivo but require light chains to trigger their release. *Immunity* 15: 105-14

- Vermeer AW, Norde W (2000) The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophys J* 78: 394-404
- Vidarsson G, Dekkers G, Rispens T (2014) IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* 5: 520
- Walsh G (2014) Biopharmaceutical benchmarks 2014. *Nature biotechnology* 32: 992-1000
- Ward PA, Antibodies NRCUCoMoPM (1999) Monoclonal antibody production. National Academies Press (US), Washington (DC)
- Wiens GD, Lekkerkerker A, Veltman I, Rittenberg MB (2001) Mutation of a single conserved residue in VH complementarity-determining region 2 results in a severe Ig secretion defect. *J Immunol* 167: 2179-86
- Williams GT, Neuberger MS (1986) Production of antibody-tagged enzymes by myeloma cells: application to DNA polymerase I Klenow fragment. *Gene* 43: 319-24
- Wilson DW, Lewis MJ, Pelham HR (1993) pH-dependent binding of KDEL to its receptor in vitro. *The Journal of biological chemistry* 268: 7465-8
- Wilson R, Allen AJ, Oliver J, Brookman JL, High S, Bulleid NJ (1995) The translocation, folding, assembly and redox-dependent degradation of secretory and membrane proteins in semi-permeabilized mammalian cells. *The Biochemical journal* 307 (Pt 3): 679-87
- Yageta S, Lauer TM, Trout BL, Honda S (2015) Conformational and Colloidal Stabilities of Isolated Constant Domains of Human Immunoglobulin G and Their Impact on Antibody Aggregation under Acidic Conditions. *Mol Pharm* 12: 1443-55
- Zalutsky MR, Narula AS (1987) A method for the radiohalogenation of proteins resulting in decreased thyroid uptake of radioiodine. *Int J Rad Appl Instrum A* 38: 1051-5